THE INTRICATE WORKINGS OF AN ENZYME FROM *NEISSERIA MENINGITIDIS* METABOLISM

A thesis submitted in partial fulfilment of the
requirements for the degree of

Doctor of Philosophy in Biochemistry

at the University of Canterbury

2015

by

Logan Colin **HEYES**

ABSTRACT

3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) catalyses the first committed step in the shikimate pathway, responsible for the biosynthesis of the aromatic amino acids as well as other aromatic metabolites. The reaction involves the aldol like condensation of erythrose 4-phosphate (E4P) and phosphoenol pyruvate (PEP) to form the product DAH7P. The presence of the shikimate pathway in plants, fungi, bacteria and apicomplexan parasites as well as its absence from mammalian cells makes enzymes of the pathway, in particular, the enzyme responsible for the first committed step, useful targets for the design of novel therapeutics. The genome of *Neisseria meningitidis* expresses a single DAH7PS, thus making it a vulnerable target for drug design. An in depth knowledge of the workings of this enzyme may provide the platform for its efficacy as a target in the fight against bacterial meningitis.

Chapter 2 reports the characterisation of three different synthetic inhibitors against *Nme*DAH7PS and illustrates which of the intermediates formed during the reaction mechanism is the better target. The identification of a single conserved active site water molecule in the active site has implications for the mechanism of *Nme*DAH7PS and the potency of the synthetic inhibitor.

Chapter 3 describes the binding and inhibition of the three aromatic amino acids at the enzyme's allosteric site. It identifies the reason(s) for the variable sensitivity and gives insight into the inactive conformation adopted by the *Nme*DAH7PS active site in the presence of an allosteric effector with respect to both substrates. It goes further to identify the effect incurred on substrate binding in the presence of allosteric inhibitor.

Chapter 4 investigates the importance of interactions at the tetramer interface for catalysis, regulation, stability and structure of *Nme*DAH7PS. Two mutations were made, one altering the only salt bridge at the Chain A-C interface and the other involving the only hydrophobic interaction at the Chain A-D interface.

Chapter 5 looks into the allosteric binding site and identified two residues that may be crucial for the binding of Phe. The substitution of these two residues and their structural and

functional characterisation illustrates the necessity for certain interactions within the largely hydrophobic binding site for binding. This chapter also carries on from work done in chapter 5 on the identification of residues important for the allosteric mechanism. Substitution of these ionisable residues and the structural and functional analysis of the resulting variant enzymes have interesting implications for the electrostatic network responsible for the transfer of allosteric signal in *Nme*DAH7PS.

This thesis gives a detailed description of the inhibitory and regulatory properties of *Nme*DAH7PS and illustrates the applicability of this work to many other systems in nature.

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 2 contains one publication reproduced with the permission of the copyright owner:

Heyes, L. C., Reichau, S., Cross, P. J., Jameson, G. B. & Parker, E. J. Structural analysis of substrate-mimicking inhibitors in complex with *Neisseria meningitidis* 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase - The importance of accommodating the active site water. *Bioorg. Chem.* **57,** 242–50 (2014). © 2014. Elsevier.

Please detail the nature and extent (%) of contribution by the candidate:

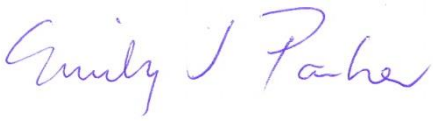
All experimental work was carried out by the candidate and the manuscript was drafted by the candidate. The candidate contributed 80% of the publication.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifies that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Emily Parker* Signature:  Date: *6 October 2015*

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 4 part I contains an accepted manuscript

Heyes, L. C.*, Cross, P. J.*, Zhang, S., Nazmi, A. R. & Parker, E. J. The functional unit of *Neisseria meningitidis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase is dimeric. *Unpublished results*

Please detail the nature and extent (%) of contribution by the candidate:

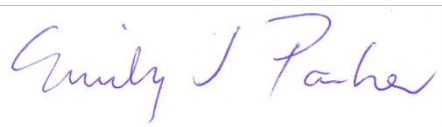
The Interface analysis, Differential scanning calorimetry, isothermal titration calorimetry, Analytical SEC, time dependent inhibition studies, crystallographic data collection and refinement (PDB code: 4UCG). The drafted manuscript was contributed to by the candidate. The candidate contributed 40% of the publication.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifies that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text



Name: *Emily Parker* Signature: Date: *6 October 2015*

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Appendix A6 contains a drafted manuscript

Heyes, L. C.*, Lang, E. J. M.*, Jameson. G. B. & Parker, E. J. pK_a variations expose allosteric communication pathways. *Unpublished results*

Please detail the nature and extent (%) of contribution by the candidate:

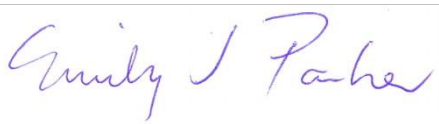
The crystallographic, soaking, data collection and refinement (PDB code: 4UC5), requiring extensive optimisation of experimental conditions, the crystal structure analysis, pH dependent isothermal titration calorimetry and pH dependent kinetics were carried out by the candidate. The drafted manuscript was contributed to by the candidate. The candidate contributed 40% of the publication.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifies that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text



Name: *Emily Parker* Signature: Date: *6 October 2015*

Table of Contents

List of figures

Number ***Page***

List of tables

Number ***Page***

Acknowledgments

What follows in this thesis is a tribute to the help and support of a variety of people.

First and foremost I would like to thank Professor Emily J Parker for all help and support she has provided me along this journey. This cannot be reiterated enough, without the help and support she has provided me, both personally and scientifically, I would not be submitting this thesis and moving on. Bearing in mind that I am only one of many in the group, the way you have an open door and ear for each and every one of us, truly is amazing. Your passion for science and desire for knowledge has rubbed off on me, for this I am grateful.

To all my friends of the Parker group, both past and present, you have made the lab, the offices and the meetings really enjoyable for me over the years. Special thanks must go to Dr Sebastian Reichau for helping me to get started in the synthetic lab, opening my eyes to the wonders of crystallography and the hollow warning, that when I entered the biochemistry lab on the 7th floor, I would probably not make it out. How right you were on that one. To Gert-Jan Moggré, the resident chemist and partner in crime, I am grateful for your discussions about "science", and providing an educated backboard for ideas that I threw your way. Your appreciation of science is really admirable and I have thoroughly enjoyed working around that vibe during my PhD. To another member of the HisG team, Dr Gerd Mittelstädt, my crystallography buddy and the only reason I could not take the Big Mac home with me. Thank you for your discussions about anything and everything, you were a great asset to all of the work undertaken by me and others in the Parker Lab. Eric Lang, I have thoroughly enjoyed our investigations into the protein that I hold dear, as well as the many tangential conversations that have arisen from these in depth discussions. To the wonderful ex-post-docs of the Parker group, Dr Penel Cross and Dr Ali Reza Nazmi for helping me with my enquiries and problems in the biochemistry lab when I was a newbie. For always being happy to discuss anything protein science related. I would like to thank all of the staff at the Australian Synchrotron MX1 and MX2 beamlines, whom I visited so frequently. Also, thank you to all of the technical staff of the department of chemistry at UC, in particular, Wayne Mackay and Matt Polson, for being

happy to help in any way possible. Thanks to Geoff Jameson, for his significant help in crystallography and his immense knowledge of “almost” everything.

I would also like to thank the Biomolecular Interaction Centre for providing me with a Doctoral scholarship and the New Zealand Synchrotron Group for the synchrotron access over the last few years.

I would like to thank my family for the love and support they have afforded me over the years.

And, last but most certainly, not least, I have a special thank you to Zoë. You have experienced this rollercoaster with me first-hand and without your ever enduring love and support I would not be in this position. I am forever grateful for the patience and understanding you have given to me over the years and I look forward to our next adventure.

A b b r e v i a t i o n s

'	adjacent chain of the tight dimer
A5P	D-arabinose 5-phosphate
ACT	aspartate kinase, chorismate mutase and tyrA
AEC	anion exchange chromatography
AUC	analytical ultracentrifugation
BTP	1,3-*bis*(*tris*(hydroxymethyl)methylamino)propane
DAH7P	3-deoxy-D-*arabino*-heptulosonate 7-phosphate
DAH7PS	3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase
E4P	D-erythrose 4-phosphate
Eco	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EPSPS	5-enol pyruvate shikimate-3-phosphate synthase
ESI-MS	electrospray ionisation mass spectrometry
G3P	glyceraldehyde 3-phosphate
G6P	D-glucose 6-phosphate
HIC	hydrophobic interaction chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
k_{cat}	turnover number
KDO8PS	3-deoxy-D-*manno*-octulosonate 8-phosphate synthase
K_M	Michaelis-Menten constant
KNF	Koshland-Némethy-Filmer
LB	lysogeny broth
MD	molecular dynamics
MWC	Monod-Wyman-Changeux
MWCO	molecular weight cut-off
Mtu	*Mycobacterium tuberculosis*
Nme	*Neisseria meningitidis*
NMR	nuclear magnetic resonance

PDB	protein data bank
PEP	phosphoenol pyruvate
PMA	phosphomolybdic acid
RMSD	root-mean-square deviation
SAR	structure-activity relationships
SAXS	small angle x-ray scattering
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SOC	super optimal broth with catabolite repression
TE	10mM tris-HCl, 0.1mM EDTA
TMAO	trimethyl-amino-*N*-oxide
Tris	tris(hydroxymethyl(aminomethane)

Amino acid single and three letter codes:

Ala	***A***	Alanine	***Leu***	***L***	Leucine
Arg	***R***	Arginine	***Lys***	***K***	Lysine
Asn	***N***	Asparagine	***Met***	***M***	Methionine
Asp	***D***	Aspartate	***Phe***	***F***	Phenylalanine
Cys	***C***	Cysteine	***Pro***	***P***	Proline
Gln	***Q***	Glutamine	***Ser***	***S***	Serine
Glu	***E***	Glutamate	***Thr***	***T***	Threonine
Gly	***G***	Glycine	***Trp***	***W***	Tryptophan
His	***H***	Histidine	***Tyr***	***Y***	Tyrosine
Ile	***I***	Isoleucine	***Val***	***V***	Valine

CHAPTER 1- INTRODUCTION

1.1 *NEISSERIA MENINGITIDIS*

Neisseria meningitidis is the bacterium responsible for meningococcal related illness. It was first isolated in 1887, however, reports of meningococcal related illness date back to the 16th century.[1] Meningococcal disease has the potential to cause major epidemics in the sub-Saharan African 'meningitis belt', and in the United States, it is the leading cause of meningitis and sepsis.[2–5] While several different bacteria can cause meningitis, *N. meningitidis* is the one with the potential to cause large epidemics. Infection with the bacteria can result in serious brain injury, hearing loss and in the absence of treatment, it is fatal in ~50% of cases.[6–8]

N. meningitidis is a Gram-negative bacterium that is exclusively pathogenic to humans and is carried asymptomatically in the human nasopharynx.[9,10] Colonisation by the bacterium begins with adhesion to epithelial surfaces, enabling invasion of the target cell.[11] Once internalised, the pathogen's cytotoxicity is mediated by tumour necrosis factor-α.[11] The mechanism of *N. meningitidis* infection begins in the mucosa where it is found dormant in ~10% of the population. Isolates of the bacteria in these mucosal surfaces have been found to be predominantly polysaccharide capsule deficient.[6] This capsule has been suggested to aid in the anti-adherent properties of the bacterium, enabling detachment from the mucosal surfaces and entry into the bloodstream.[9] Isolates of the bacteria from the blood stream and cerebrospinal fluid are predominantly polysaccharide encapsulated and because of this protective capsule the bacteria is a poor antigen, has serum resistance and can go undetected by the immune system.[12,13]

The septicaemia associated with the disease, only occurs once the disease has traversed the mucosal barrier into the host's blood stream. Upon crossing the blood-brain barrier, it causes serious infection of the meninges that line the skull and vertebral canal, which encloses the brain and spinal cord.[7]

The disease itself is categorised into twelve Serogroups, 6 of which have been known to cause epidemics (A, B, C, W-135, X and Y).[8] Serogroups, A and C are the most predominant throughout Asia and Africa while B and C are the most prevalent in countries of the developed world including the Americas, Western Europe and Australia.[7,14] In New Zealand an outbreak

of Serogroup B disease was experienced in 2004 and the resulting campaign involving the outer membrane vesicle (OMV) vaccine has since contributed to its decreasing prevalence.[15,16]

Routine vaccination against the disease remains the most effective strategy for controlling the disease. There are several vaccines available to control the disease: a meningococcal A conjugate vaccine, C conjugate vaccines, tetravalent A, C, Y and W-135 conjugate vaccines and meningococcal polysaccharide vaccines.[8] However, polysaccharide vaccines for serogroup B are challenging due to antigenic mimicry of components of human neurologic tissues by the polysaccharide capsule, thus making the immune response very poor.[8,14] The first vaccine (4CMenB) against the serogroup B strain (MC58), consisted of 4 components and was released for administration in 2014.[8] The vaccine was composed of two recombinant fusion proteins (*Neisseria* heparin-binding antigen and factor H-binding protein), recombinant Neisserial adhesion A and detergent-treated OMV's.[14]

Vaccines remain a preventative measure and upon infection, treatment of the disease with broad spectrum antibiotics such as chloramphenicol, ampicillin, penicillin and ceftriaxone is the most effective strategy.[17] However, this is dangerous, as the ability for the *N. meningitidis* to develop antibiotic resistance is an increasing problem.[17] Therefore, the necessity for novel antibiotics against particularly the Serogroup B strain (MC58) of *N. meningitidis* is becoming increasingly important.

1.2 THE SHIKIMATE PATHWAY

The presence of the shikimate pathway in plants, fungi, bacteria and apicomplexan parasites as well as its absence from mammalian cells makes it a viable target for drug design and discovery.[18,19] This pathway is responsible for the biosynthesis of the aromatic amino acids, L-phenylalanine (Phe), L-tyrosine (Tyr) and L-tryptophan (Trp) as well as many other aromatic metabolites (Figure 1.2.1). The first committed step of the pathway is catalysed by 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) and involves the aldol-like condensation of erythrose 4-phosphate (E4P) and phosphoenol pyruvate (PEP) resulting in the seven carbon sugar, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAH7P). The formation of chorismate requires a further six enzymatic steps. At this point the pathway diverges via anthranilate synthase to form the precursor for tryptophan (anthranilate) and via chorismate mutase to form the precursor for phenylalanine and tyrosine (prephenate), or to generate precursors for folate (*p*-aminobenzoic acid) or vitamin K (isochorismate) biosynthesis.[18] Unlike plants and bacteria that rely on the shikimate pathway for synthesis of their aromatic amino acids, mammals are able to obtain these essential metabolites from their diet. The flux of the pathway is regulated at the protein level via DAH7PS, making it, the first committed step in the pathway an ideal target for the design and discovery of novel antibiotics.[20,21,22]. This pathway has already been used utilised as a target in herbicides. For instance, Glyphosate, a commonly used herbicide is an inhibitor of 5-enol pyruvate shikimate-3-phosphate synthase (EPSPS).[23] DAH7PS is a curious enzyme, as it uses a variety of complex methods to achieve allosteric regulation, whether it be through protein complex formation, gene fusion, insertions or extensions to achieve simple feedback inhibition.[20,21,24–29]

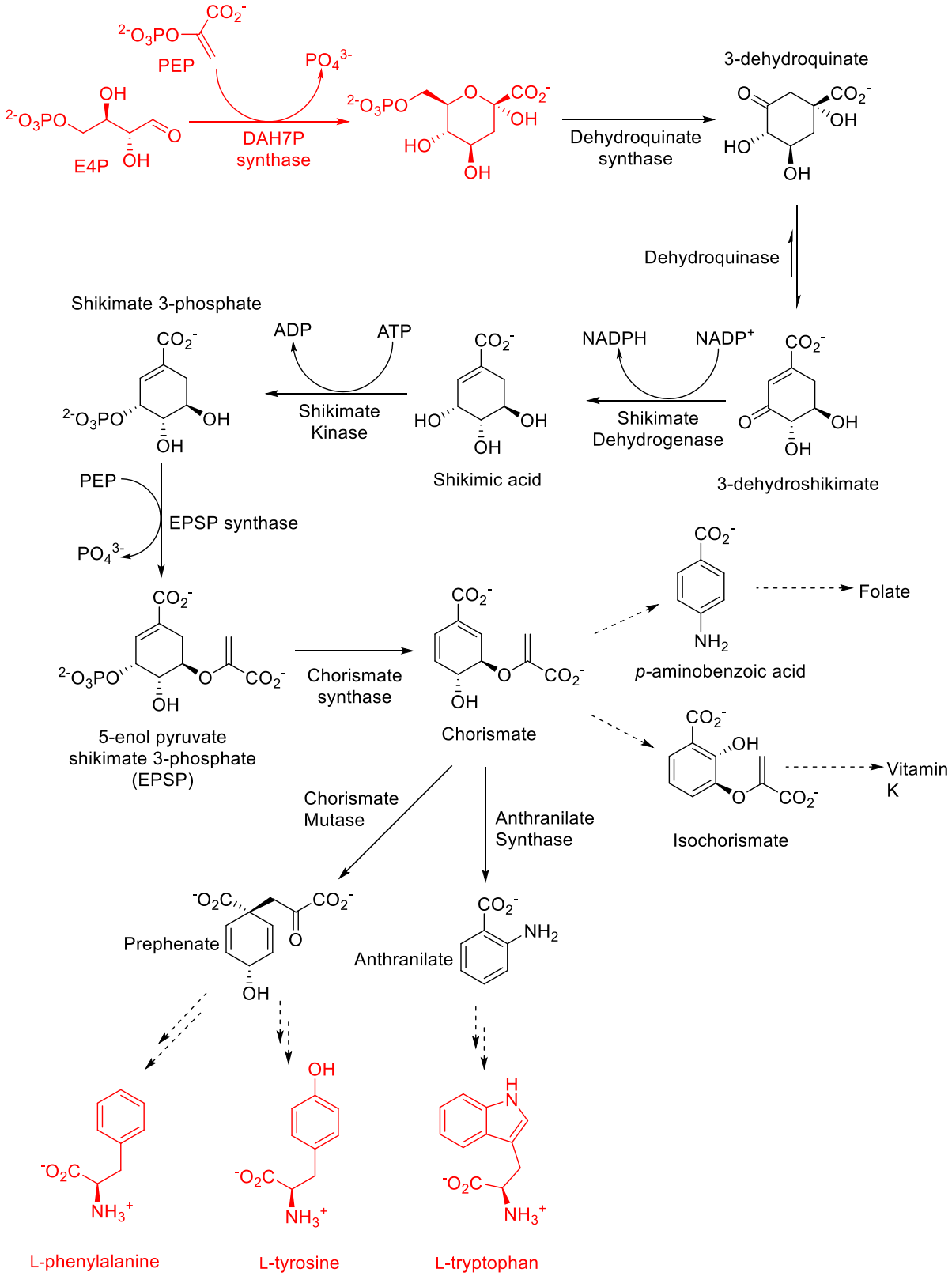


Figure 1.2.1 The shikimate pathway illustrating the branch point chorismate and its multiple end-products. Highlighted in red is the first committed step catalysed by DAH7PS and the aromatic amino acid end-products; L-Phe, L-Tyr and L-Trp.

1.3 3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE (DAH7PS)

1.3.1 Classification of DAH7PS

DAH7PS can be classified into two distinct families based on sequence identity and molecular weight, denoted type I and type II. Type I DAH7PS enzymes can be further divided into two subfamilies, denoted type Iα and type Iβ based purely on sequence identity. The classification of the DAH7PS subfamilies is centred on these modifications to the core barrel that tend to be responsible for allostery (Figure 1.3.1.1).[20]

DAH7PS from the type Iα subfamily like those from *Escherichia coli*[22,30,31], *Saccharomyces cerevisiae*[32–34] and *N. meningitidis*[26] are all below 40 kDa in size, whereas their type II orthologues are all larger, like that of *Mycobacterium tuberculosis* (~54 kDa).[35] Type Iβ DAH7PSs like those from *Pyrococcus furiosus*[36] *and Thermotoga maritima*[37], are smaller (~30-38 kDa) and share little sequence identity with their type Iα counterparts. Members of the type Iβ subfamily are more closely related to the enzyme 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase (KDO8PS), an enzyme involved in the biosynthesis of the cell wall of Gram-negative bacteria, which has the ability to accommodate D-arabinose 5-phosphate (A5P) as a substrate instead of E4P.[38–42] A limited number of organisms have been shown to contain both type I and type II DAH7PS, including, *Pseudomonas aeruginosa*[43], *Xanthomonas campestris*[44], *Stigmatella aurantiaca*[45] and *Amycolatopsis methanolica*[46]. The type II DAH7PS from these organisms tends to be involved in secondary metabolism.

A point of comparison between the different subfamilies of DAH7PS comes about from the extensions and modifications to the core $(\beta/\alpha)_8$ barrel, which is the structural fold for all known DAH7PS proteins (Figure 1.3.1.1). Type Iα DAH7PS enzymes typically contain a double β-sheet insertion and double α-helix N-terminal extension. Type Iβ enzymes are either not regulated (*Pfu*DAH7PS)[36], or have a large N-terminal extension that is reminiscent of either a chorismate mutase (*Gsp*DAH7PS)[47] or aspartate kinase, chorismate mutase and tyrA (ACT) (*Tma*DAH7PS)[25] domain. Type II enzymes characteristically have a larger insertion of two α-helices, and a very large N-terminal extension.[48] These extensions, insertions and accessory

domains are believed to be intrinsically associated with the regulation of these enzymes by the pathway end-products.[20]

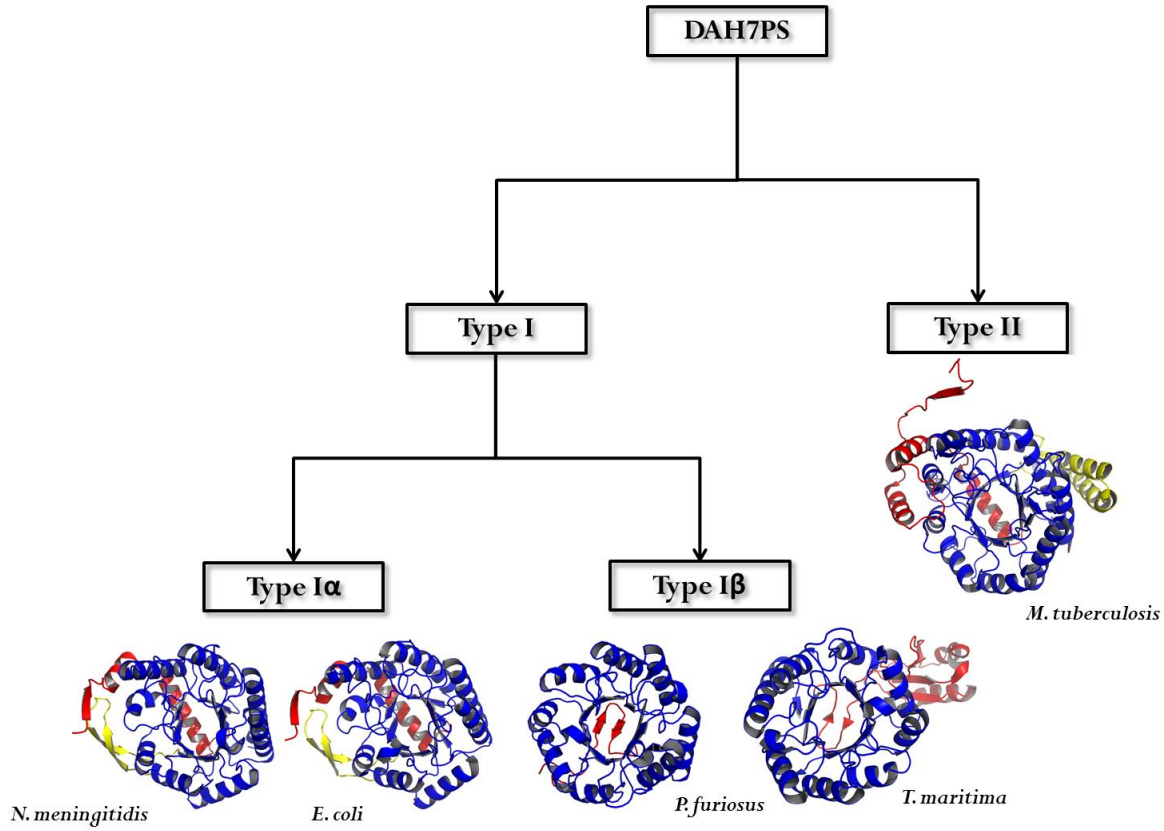


Figure 1.3.1.1 Classification of DAH7PS into their appropriate subfamilies. Core $(\beta/\alpha)_8$ barrel coloured in blue. B-sheet or α-helix insertions are coloured in yellow. N-terminal extensions are coloured in red. Type Iα DAH7PS represented by N. meningitides (PDB code:4HSN) and E. coli (1KFL). Type Iβ DAH7PS represented by P. furiosus (PDB code:1ZCO) and T. maritima (PDB code: 1RZM). Type II DAH7PS represented by M. tuberculosis (PDB code: 2B7O).

1.3.2 Mechanism of DAH7PS

DAH7PS catalyses the aldol-like condensation between E4P and PEP, generating the seven carbon sugar DAH7P. Many of the key mechanistic details of the reaction have been elucidated via labelling[32], alternative substrate,[40,49] structural[28,31,33,35,50] and inhibitor studies.[51–54] The enzyme catalysed reaction is believed to proceed in an ordered sequential fashion.[55–57] The mechanism is stereospecific with respect to both substrates, where the *si* face of PEP nucleophilically attacks the *re* face of E4P (Figure 1.3.2.1).[58–61]

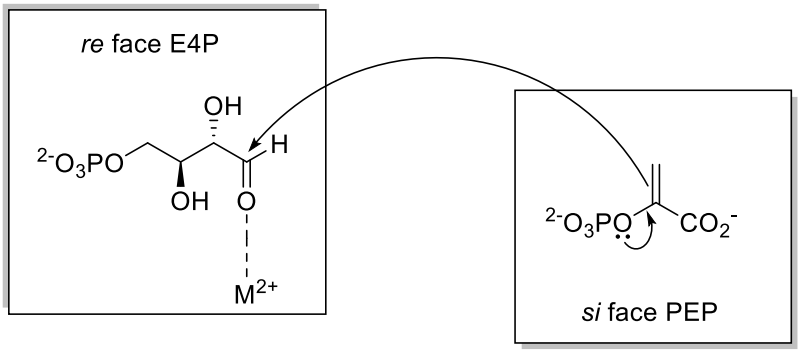


Figure 1.3.2.1 Diagram illustrating the facial selectivity of the nucleophilic attack of the allylic carbon of PEP on the aldehyde moiety of E4P, which is activated by coordination to the divalent active-site metal ion.

The aldol-like reaction begins with the cleavage of the C-O bond to the phosphate group of PEP. Activation of the aldehyde functionality of E4P as an electrophile by the essential active-site divalent metal ion enables the nucleophilic attack of the enol moiety of PEP (Figure 1.3.2.2). The resulting oxyanion is protonated by an active site Lys residue (part of a conserved KPR(T/S) motif). Formation of a bond between C_1 of E4P and C_3 of PEP yields a transient 7-carbon oxocarbenium ion intermediate.[54] The stereospecific nucleophilic attack of a conserved active site water molecule, which is activated via deprotonation by an active site residue, results in a phosphohemiketal tetrahedral intermediate. Upon elimination of phosphate, the product loses its stereochemistry at C_2, giving the acyclic DAH7P. This acyclic form of DAH7P is in equilibrium with its pyranose form.

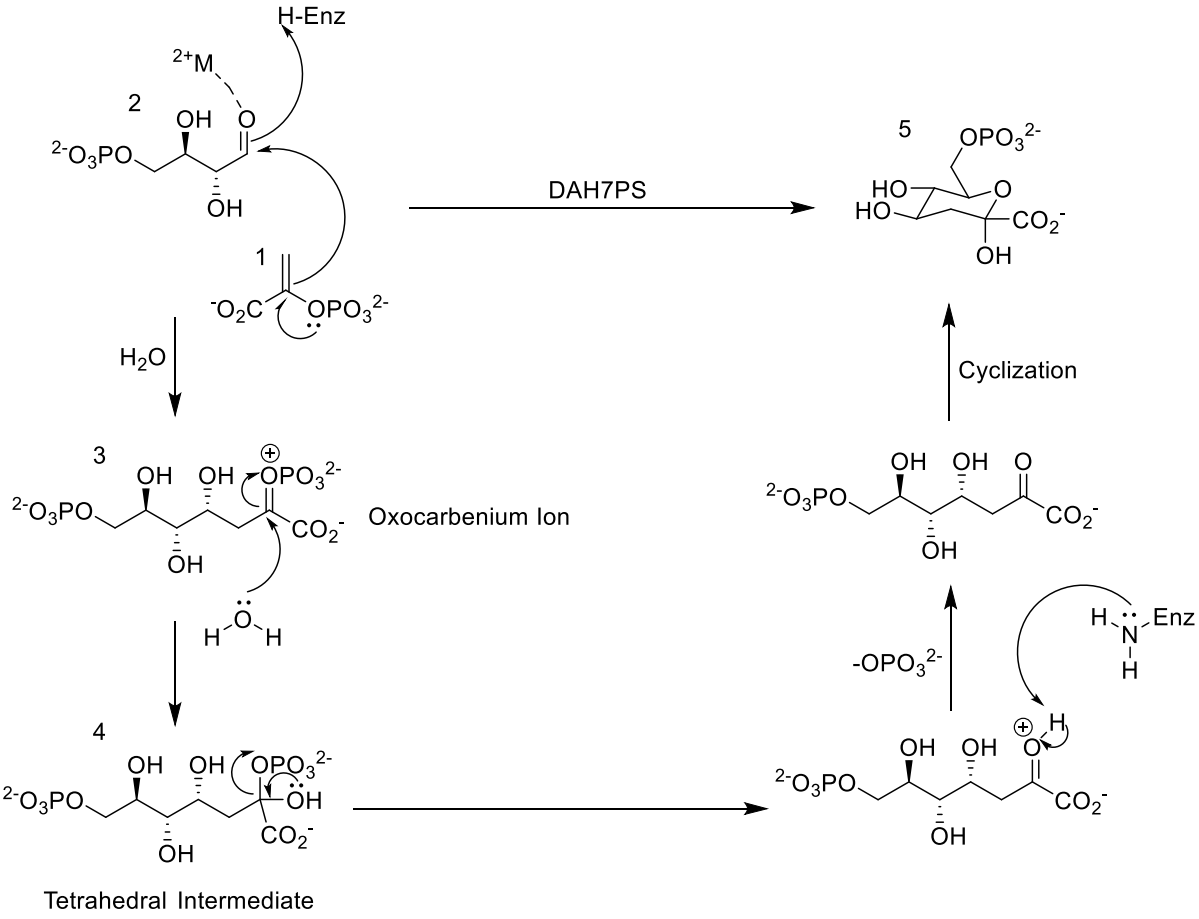


Figure 1.3.2.2 Mechanism of the DAH7PS catalysed reaction, illustrating its procession via the oxocarbenium and tetrahedral intermediate. 1, phosphoenol-pyruvate; 2, D-erythrose 4-phosphate; 3, Oxocarbenium ion intermediate; 4, Tetrahedral intermediate; 5, DAH7PS.

1.3.3 The role of the active site water

In the mechanism of the DAH7PS catalysed reaction, the oxocarbenium ion is a transient species, which is rapidly attacked by water, resulting in the labile phosphate hemiketal intermediate. With the influx of structural information on this enzyme with substrates,[26,31,33,35,40,50,62,63] substrate mimics[63] and intermediate mimics[54,64] bound at the active site, the water molecule involved in the reaction is hypothesised to sit very close to C_2 of PEP (Figure 1.3.3.1). Previous theories of the enzyme-catalysed reaction mechanism from the related KDO8PS enzyme have proposed that nucleophilic attack of water occurs at C_2 of PEP, which is made possible via coordination to the active site metal ion (carbanion mechanism).[41,65,66] This mechanism unfortunately does not make any structural or chemical sense, as it would require PEP to possess an unprecedented electrophilic character at C_2 and

upon nucleophilic attack of an active site water molecule, the resulting carbanion would have a pK_a estimated to be >30. Due to the presence of carbon-hydrogen bonds of greater acidity in addition to water molecules, carboxylic acids and ammonium ions in the enzymes active site, the carbanion formed would likely be quenched before being able to attack E4P.[67] Therefore, it is likely that the formation of an oxocarbenium ion intermediate via a Lewis-acid type mechanism is the most probable.

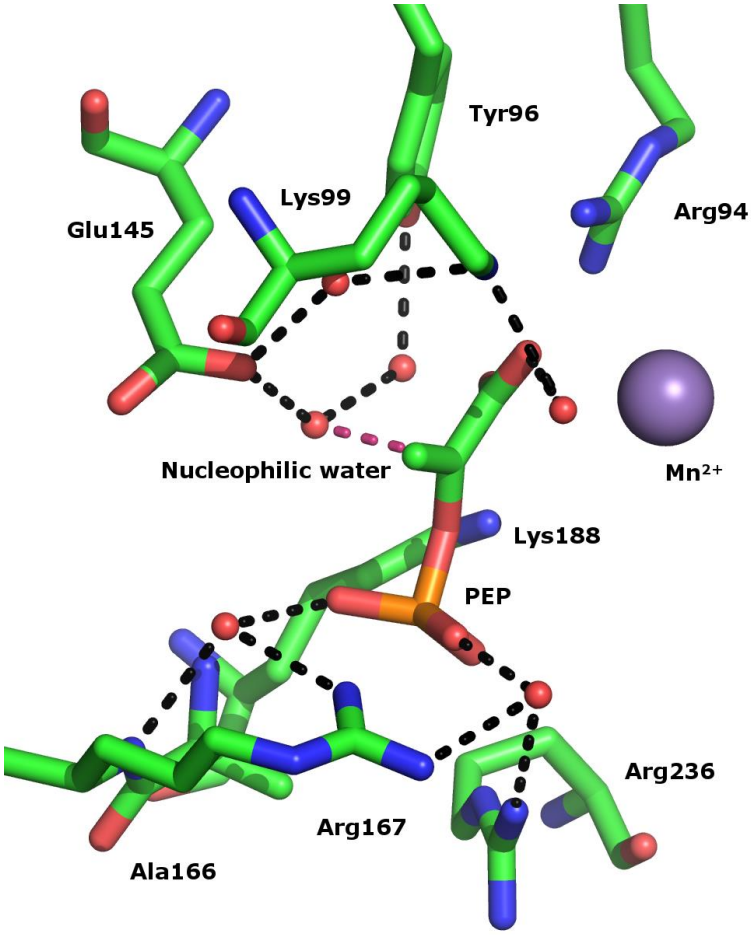


Figure 1.3.3.1 Active site of NmeDAH7PS illustrating complex network of waters present during PEP presence. Hydrogen bonds involving water molecules indicated by black dashed lines. Proximity of nucleophilic water to C_2 of PEP is indicated by a maroon dashed line.

The stereochemistry of the tetrahedral hemiketal intermediate is dictated by the facial selectivity of the nucleophilic attack of the water at C_2 the oxocarbenium ion. Given the recent work on enantiopure tetrahedral intermediate mimics and their relative potency, we can hypothesise that this attack occurs on the *re* face of the oxocarbenium ion.[51,54,64] There are four conserved water molecules that are present both the PEP-bound *Nme*DAH7PS structure (PDB code 4HSN) and the tetrahedral intermediate mimic-bound *Mtu*DAH7PS structure (PDB

code 3PFP).[26,54] Three of these are located on the *re* face of these bound ligands and only one is close enough to interact at the C_2 position.

It is has been described that the proficiency of a synthetic inhibitor is due to its ability to mimic the most reactive species during the catalytic mechanism.[68] Trapping the enzyme in a state where it is stabilising the most reactive species has been the best method of obtaining irreversible-like dissociation constants.[69,70] The ability of a synthetic inhibitor to mimic either the tetrahedral or the planar oxocarbenium ion intermediates whilst still appropriately accommodating the nucleophilic water molecule, is predicted to result in a highly potent inhibitor of DAH7PS.[54]

1.3.4 Structure of DAH7PS

DAH7PS enzymes all share the same core barrel structure, despite the low sequence similarity between enzymes from different families. The structural characterisation of these enzymes by X-ray crystallography has revealed that they share a $(\beta/\alpha)_8$ barrel fold, which houses the active site machinery. The enzymes of different types vary in their extensions and/or insertions to the barrel and in their quaternary structure. These modifications have been shown to play an important part in the regulation of DAH7PS.[20]

Type Iα enzymes such as those from *E. coli, S. cerevisiae* and *N. meningitidis* contain modifications to the core barrel structure that are believed to be involved in the enzymes' allosteric response.[22,26,33] A double β-sheet insertion forms the bulk of the allosteric binding site and an N-terminal extension closes over the allosteric ligand once bound. These enzymes consist of four identical monomeric units that arrange themselves as dimers of dimers forming the homotetrameric assembly. The lack of a discrete regulatory domain in these enzymes makes their regulatory mechanisms intriguing.

Type Iβ enzymes have acquired allostery through different means. They have recruited large domains for their regulation that are attached at either the N or C terminus. The DAH7PSs from *Geobacillus sp* and *Listeria monocytogenes* have a dimeric chorismate mutase-like domain, which binds and catalyses the rearrangement of chorismate to form prephenate.[21,71] It is believed that regulation of DAH7PS activity occurs through adoption of a conformation in

the presence of prephenate that limits the activity of DAH7PS.[47] *Tma*DAH7PS contains an N-terminal extension with structural and sequence homology to the ACT domain. Crystal structures and small angle X-ray scattering data have shown that this ACT-like domain undergoes a large conformational shift to gate the active site in the presence of tyrosine.[25] The truncated form of this enzyme that lacks the ACT-like regulatory domain loses sensitivity to tyrosine, much like that of the unregulated *P. furiosus* DAH7PS also from the type Iβ subfamily.[36] Interestingly, fusion of the ACT domain from *T. maritima* onto the unregulated *P. furiosus* DAH7PS resulted in a chimeric enzyme sensitive to inhibition by tyrosine, illustrating that allostery can be transferred between members of the type Iβ subfamily.[72]

To date the only type II DAH7PS to be structurally characterised is from *M. tuberculosis.*[48,73] The structure of this enzyme contains the same core catalytic $(\beta/\alpha)_8$ barrel as the type I DAH7PSs, but the primary reason for its ~10% sequence identity with its type I counterparts is the extensive modifications to the core structure. These modifications make up three different binding sites, enabling regulation by all aromatic amino acids in a synergistic manner.[24,27,48]

1.4 ALLOSTERY

Allostery is an important characteristic of every living cell and because of this it is becoming ever more important in drug discovery efforts. Allostery involves the non-covalent binding of an effector molecule to a protein, aiding in the regulation, or activation of that protein. This process occurs through the binding of an effector molecule at a remote (allosteric) site, resulting in a change in, structure, dynamics or electrostatics at the active site.[74] Allostery is a feature that is present throughout the five kingdoms and its purpose is to regulate the activity of important cellular processes.

Enzymes have a tendency to form multimeric structures and associate as homo-mers, usually consisting of even numbers of subunits, generating symmetry related interfaces.[75] They also have been known to form complexes with other enzymes in the pathway. These two phenomena give rise to multiple binding sites and the possibility of cooperativity between these sites.[76] In these cases, cooperativity (positive and negative) between binding sites can enable precise control on cellular demand for metabolites. To date there have been a few major models that have contributed to our understanding of allostery.[77]

1.4.1 Models for allostery in proteins

The Monod-Wyman-Changeux (MWC) model, states that in the absence of effector molecule, the protein exists in equilibrium between two conformational states; relaxed (R) and the tense (T) states.[78] The binding of an effector molecule shifts this equilibrium to favour the state that has a higher affinity for the effector molecule by stabilising a global conformational change in the protein's quaternary structure (Figure 1.4.1A).

The Koshland-Némethy-Filmer (KNF) model, states that the binding of an effector molecule causes a change in the structure of the protein, enabling better accommodation of the effector molecule.[79] The binding of an effector molecule to a particular subunit of the protein induces a change in that subunit, encouraging or discouraging the binding of the second effector molecule to that subunit. This model describes a progressive change in conformation resulting in an induced fit like mechanism for binding several allosteric effectors.

It does however, allow the relative subunits to adopt different conformations, therefore giving rise to the notion of positive and negative cooperativity, where the affinity of the protein for the same ligand is altered based on the presence or absence of that ligand (Figure 1.4.1B).

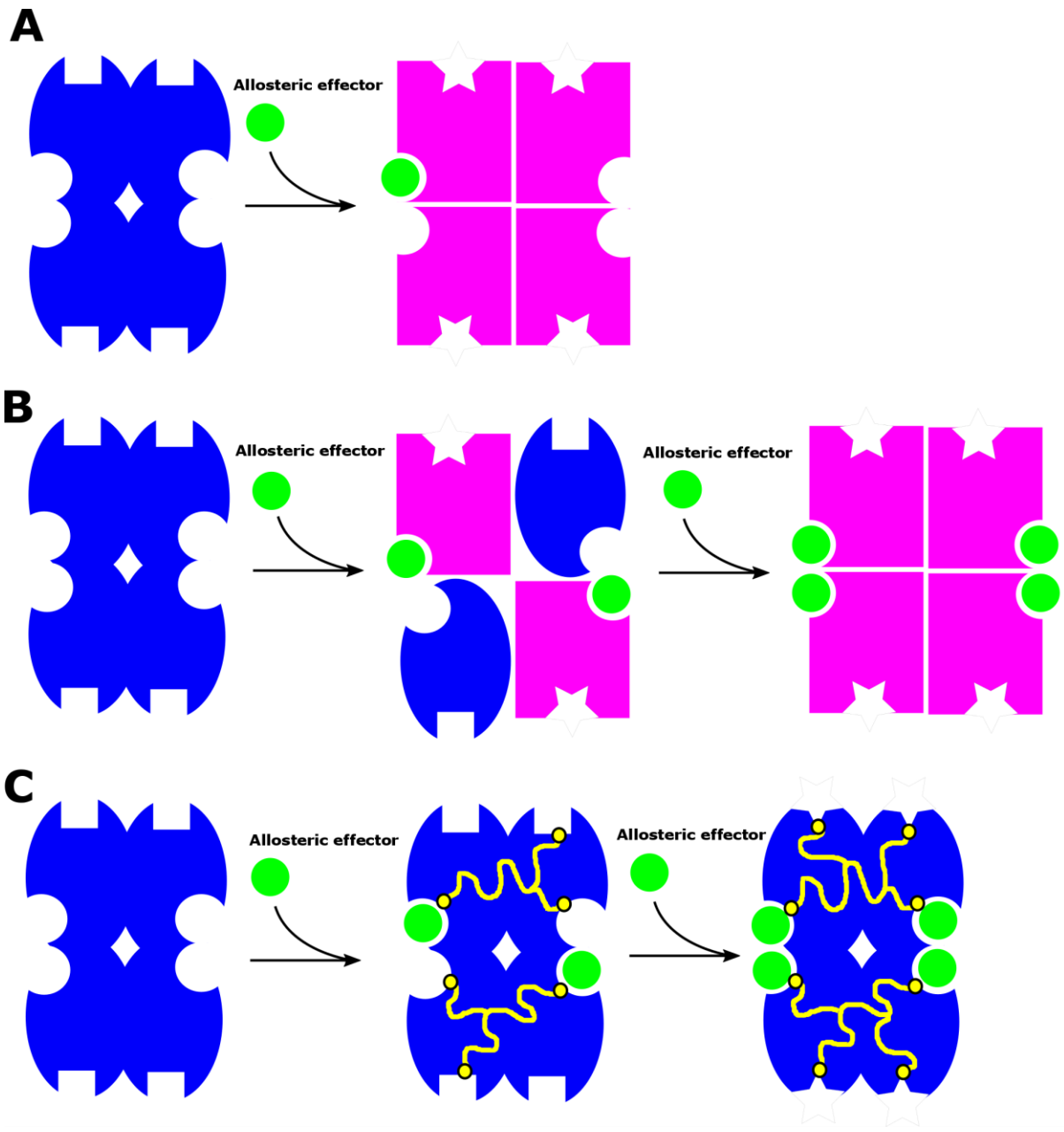


Figure 1.4.1 Models for allosteric regulation. A. Monod-Wyman-Changeaux model where the protein adopts either the relaxed 'R' state or the tense 'T' state. B. Koshland-Némethy-Filmer model where ligand binding causes a conformational change that facilitates binding of another ligand. C. Allostery without conformational change. Binding of a ligand propagates a signal that is registered at the other allosteric site and the active site.

There are obviously limitations to the aforementioned models of allostery, as they both describe conformational change. Recently there have been developments in the modelling of allostery in a more broadly applicable manner that incorporates changes to proteins in other

ways, elicited by ligand binding.[80] Proteins that undergo no conformational change during allostery or are intrinsically unstructured are exceptions to the MWC and KNF models.[81] Such exceptions gave rise to the dynamic energy landscape model and the notion that proteins exist as an ensemble of possible conformational states, which are populated according to their Boltzmann distribution.[82] This ensemble model of protein dynamics, enables sampling of a large energy landscape with multiple energy minima, favouring the lowest energy conformation across the landscape. The binding of an effector molecule in this model restructures the energy landscape, redistributing the protein population, favouring new energy minima. Allostery therefore involves far more than the conformational transition from 'R' to 'T' state but rather a dynamic equilibrium between sub-states of the energy landscape.[83]

The aforementioned traditional views of allostery are governed by a kinetic concept, where the contraction and release of strain caused by ligand binding governs the allosteric effect at the active site via a single pathway.[84] However, more contemporary views are founded on the nature of thermodynamic equilibria, which are based on population shift during allostery, giving rise to the propagation of multiple signals through complex networks of interactions (Figure 1.4.1C).[80,82,85–90] Population shifts upon allosteric ligand binding with an accompanied conformational change could be dominated by enthalpy or entropy. Variation in free energy without conformational change, tend to be governed by changes in entropy. Conversely, in enthalpically driven allostery, all propagation pathways effectively culminate in the required conformational change.[81] In systems that undergo allostery without conformational change, treatment of the protein as a dynamic, constantly changing entity that allows the propagation of multiple signals from the effector binding site is paramount. These signals form communication pathways, which alter the dynamic, structural and electrostatic properties of remote sites enabling cooperativity and regulation.[87,91]

1.5 NEISSERIA MENINGITIDIS DAH7PS

1.5.1 Introduction

The genome of *N. meningitidis* encodes a single DAH7PS, which is primarily regulated by the pathway end-product Phe, but is sensitive to inhibition by both Trp and Tyr at high concentrations.[26] Comparison of the *Nme*DAH7PS sequence with DAH7PSs from other organisms indicates that it belongs to the type Iα subfamily.

1.5.2 Structure

The monomeric subunit of *Nme*DAH7PS contains a core $(\beta/\alpha)_8$ barrel fold, which is consistent with all other structurally characterised DAH7PS enzymes. There is a single N-terminal α-helix and loop extension to the core barrel as well as a double β-sheet insertion (Figure 1.5.2.1A). At the N-terminus a tail (residues 1-11) is followed by a β-strand (β0, residues 12-16) and then the β0α0 loop (residues 17-20). This loop leads into the α0 helix (residues. 21-28), α0α00 loop (residues 29-31) and the large α00 helix (residues 32-50). This N-terminal extension is quite large and packs against the outside of the core barrel. The β-hairpin insertion consists of residues 209-224, and is positioned on the outside of the monomeric unit. These modifications contribute significantly to the inter-subunit interfaces and are similar to that of all structurally characterised DAH7PS enzymes in that they play a pivotal role in the creation of the allosteric binding site.

The biological unit of *Nme*DAH7PS is homotetrameric, consisting of two dimers, one of which comprises an extensive interface (denoted tight dimer interface) and the other a far weaker interface (termed the tetramer interface) (Figure 1.5.2.1B). The tight dimeric units may contribute to the method adopted by *Nme*DAH7PS to achieve regulation. The β-hairpin insertion contributes to the tight dimer interface and the N-terminal extension wraps around the base of the tight dimer at the tetramer interface, interacting with adjacent subunits to aid in stabilisation of both the dimeric and tetrameric *Nme*DAH7PS.

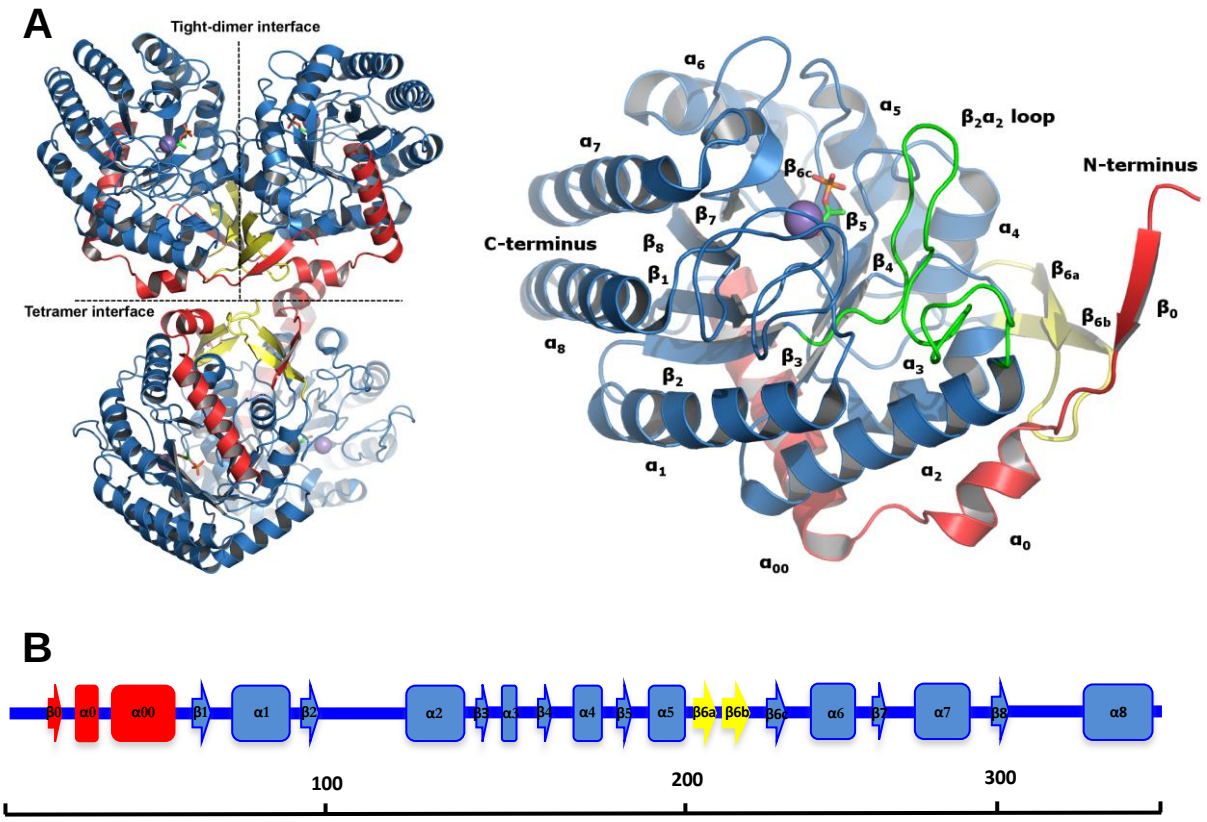


Figure 1.5.2.1 A) NmeDAH7PS (PDB code 4HSN), coloured by domain extension (red), core catalytic unit (blue), insertion (yellow), $\beta_2\alpha_2$ loop (green), active site metal ion Mn^{2+} (purple) and PEP-bound at the active site (green sticks) Top left. Tetrameric assembly of NmeDAH7PS. Top right. Monomeric unit of NmeDAH7PS with labelled secondary structural features. B) Schematic of the position of all secondary structural features in chain.

1.5.3 Active site of *Nme*DAH7PS

The active site of *Nme*DAH7PS resembles that of all other known DAH7PSs. While the sequence similarity across different families and sub-families is low, there are several residues localised at the active site that remain absolutely conserved, resulting in remarkably similar active site architectures.

1.5.3.1 Metal binding site

The metal binding site of *Nme*DAH7PS is similar to all other structurally characterised DAH7PS enzymes. From the PEP-bound crystal structure (PDB code: 4HSN) it is evident that there are four residues (Cys63, His270, Glu304 and Asp324) that coordinate the preferred Mn^{2+} metal ion in a trigonal bypyrimidal fashion (Figure 1.5.3.1.1).[26] This leaves one coordination site

available for activation of the aldehyde of E4P upon binding. However, in the PEP-bound *Nme*DAH7PS structure, Mn^{2+} coordinates a water molecule in the absence of E4P.

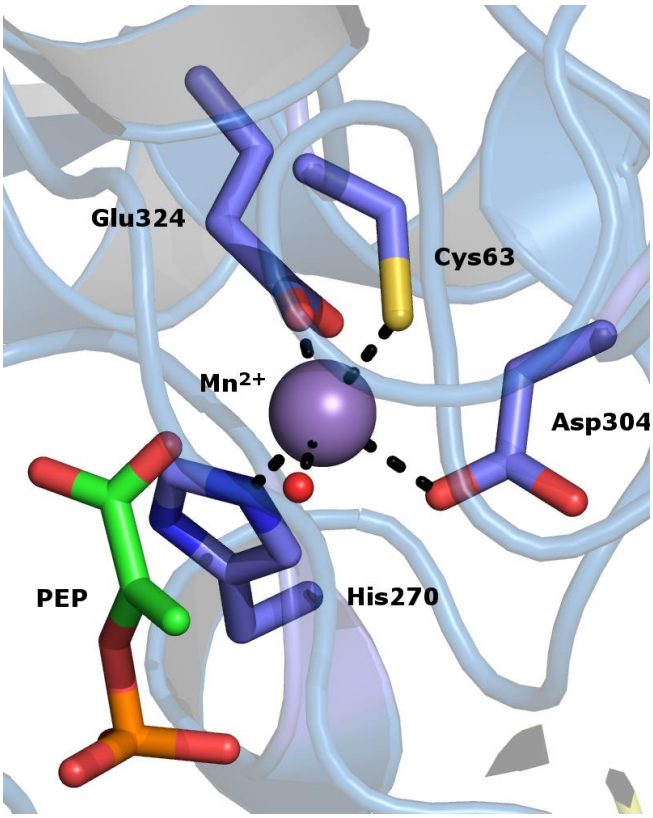


Figure 1.5.3.1.1 Metal binding site of NmeDAH7PS coloured in blue (PDB code 4HSN) . Manganese shown as purple sphere. Coordinating atoms denoted by black dashed lines

1.5.3.2 PEP BINDING SITE

In the PEP-bound *Nme*DAH7PS crystal structure (PDB code: 4HSN), the highly negatively charged PEP interacts with positively charged side-chains and does not directly coordinate to the active site metal ion, which is in agreement with previous investigations.[26,92] The phosphate group of PEP is held securely in place via salt bridges with Arg167, Lys188, Arg236 and a hydrogen bond with the backbone peptide amino group of Ala166 (Figure 1.5.3.2.1). The linking oxygen of the phosphate group of PEP forms a hydrogen bond with Lys188. The carboxylate of PEP is held in place via salt bridges with Arg94 and Lys99. These residues (Arg94, Lys99, Arg167 and Arg236), which interact directly with PEP, are conserved across all DAH7PS subfamilies. There is a network of water molecules present in the active site, held in place by the substrate PEP itself, or by residues Tyr96 and Glu145, which are hypothesised to

play a crucial role in the nucleophilic attack of the oxocarbenium ion upon its formation during catalysis.

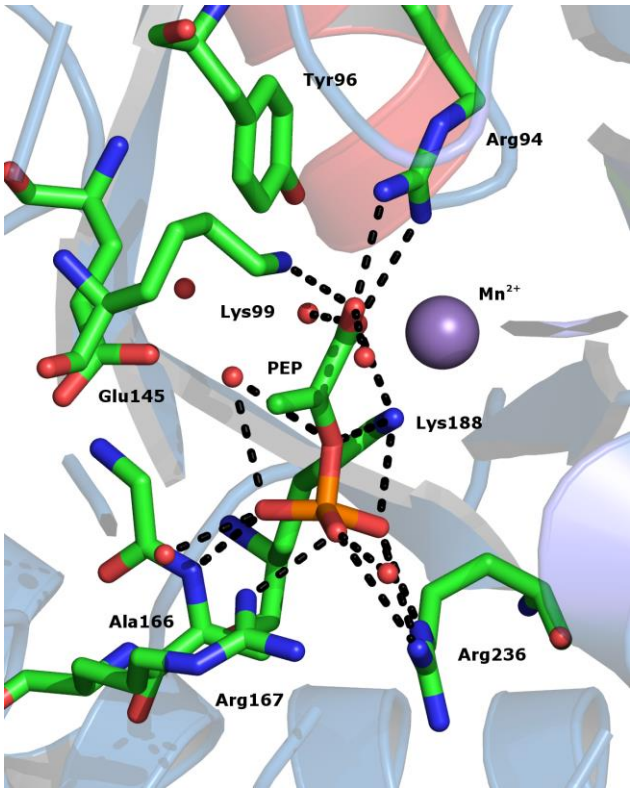


Figure 1.5.3.2.1 The binding site of PEP (PDB code 4HSN). Key interactions between PEP and residues of the active site are shown with black dashed lines.

1.5.2.3 E4P BINDING SITE

The binding site for E4P is more difficult to elucidate, as direct crystallographic information of E4P binding is difficult to obtain. This is due to the likelihood of catalytic turnover upon soaking E4P into the PEP-bound crystal. Interestingly there is a conserved (KPR(T/S)) motif that is found across all DAH7PS enzymes. This motif is believed to responsible for the binding and orientation of E4P in the active site for reaction with PEP. This motif is positioned on the highly flexible $\beta_2\alpha_2$ loop of *Nme*DAH7PS.

There have been two occasions where the binding of E4P has been identified experimentally. The first involved the soak of E4P into a crystal of thermophilic type Iβ *Tma*DAH7PS at 4°C. This cold soak of a thermophilic enzyme enabled identification of both substrates in the

crystal structure due to the limited turnover at this temperature (PDB code 1RZM).[31] Unfortunately, E4P adopted an unlikely binding mode in this crystal structure. However, in conjunction with the information obtained from the crystal structure of *S. cerevisiae* DAH7PS (PDB code 1OF8), which was soaked with an analogue of E4P, glyceraldehyde 3-phosphate (G3P), the binding mode of E4P could be properly modelled.[33] The valuable information obtained from the positioning of the phosphate and hydroxyl groups in these structures helped guide the modelling, design and synthesis of several potent tetrahedral intermediate mimics as inhibitors of *Mtu*DAH7PS.[54] The inability for the active site of *Mtu*DAH7PS to bind smaller or larger intermediate mimics with the same affinity, illustrates the selectivity of the site for a seven carbon mimic.[64] From the computational modelling experiments and the conservation of residues in the active site of DAH7PS, it is likely that in *Nme*DAH7PS a hydroxyl group of E4P interacts with the proline backbone carbonyl group of the KPR(T/S) motif. Lys99 is proposed to be involved in the protonation of the aldehyde of E4P during reaction with PEP. Arg101 is predicted to be involved in binding the phosphate group of E4P. All of these residues are part of the highly conserved KPR(T/S) motif and therefore give credence to the proposed binding mode of the natural substrate E4P.

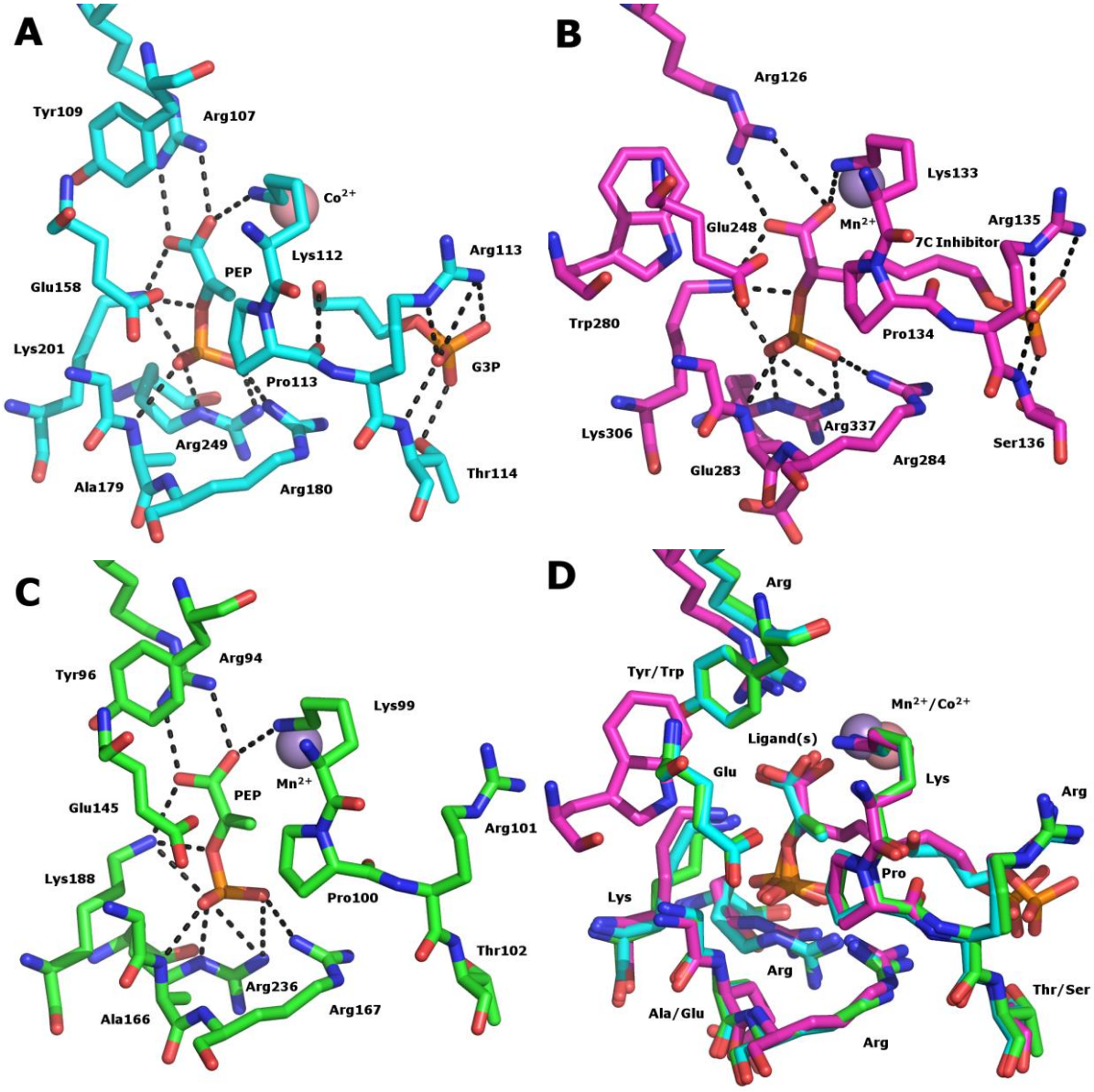


Figure 1.5.3.3.1 Comparison of E4P binding sites from S. cerevisiae, M. tuberculosis and N. meningitidis. Important interactions are illustrated with black dashed lines. A. SceDAH7PS E4P binding site showing important residues and ligands bound to enzyme (PDB code 1OF8). B. MtuDAH7PS E4P binding site illustrating important residues for tetrahedral intermediate binding (PDB code 3PFP). C. NmeDAH7PS E4P binding site with only PEP-bound (PDB code 4HSN). D. Overlay of all three structures indicating both sequence and structural conservation of PEP and E4P binding site.

Comparison of these different structures (PDB codes: 3PFP, 1OF8 and 4HSN) reveal some interesting features of the E4P binding site in DAH7PS (Figure 1.5.3.3.1). Firstly, there is a high degree of sequence conservation in this binding site, however, in *Eco*DAH7PS and *Nme*DAH7PS, a tyrosine residue is responsible for positioning of the nucleophilic water molecule, whereas in *Mtu,* this is replaced with a tryptophan.[54] The backbone peptide amino group of the alanine found in *Sce*DAH7PS and *Nme*DAH7PS forms a hydrogen bond with the

phosphate of PEP, however in *Mtu*DAH7PS, this residue is replaced with a glutamate. This glutamate residue in *Mtu*DAH7PS, may play important roles in mediating an allosteric signal from one of its three functionally distinct binding sites.[24] The only other variation is in the KPR(T/S) motif, where, for *Mtu*DAH7PS, the threonine is replaced with a serine. This residue is functionally the same across these organisms. Overall, the binding site for type I and type II DAH7PS is structurally similar and functionally identical. This enables confident translation of active site characteristics between different organisms.

1.5.4 Allostery in *Nme*DAH7PS

*Nme*DAH7PS does not require complex formation like that seen in the type II *Mtu*DAH7PS[93], nor does it have a large discrete domains for regulation, like that seen in type Iβ DAH7PS.[21,25] All structures of type Iα DAH7PS with an allosteric ligand bound show no major conformational change upon ligand binding, therefore the mechanism of allostery in this enzyme is intriguing.

Kinetic analysis of the inhibitory properties of the pathway end-products (Phe, Tyr and Trp) has revealed that *Nme*DAH7PS is sensitive to all three aromatic amino acids at high concentrations, but is most sensitive to inhibition by Phe.[26] Isothermal titration calorimetry (ITC) experiments have shown that there is cooperativity between allosteric sites, obtained from the binding isotherm fit to a sequential two-site binding model.[26]

The allosteric binding site in *Nme*DAH7PS is formed by the modifications to the core catalytic barrel (PDB code: 4HSO). The N-terminal extension closes over the site and provides a scaffold to aid in closure of the site (α00-helix). The double β-strand insertion forms a large part of the binding site and plays a large role in ligand binding (Figure 1.5.4.1).

Inspection of the binding site of the Phe- and Tyr-bound structures from *Eco*DAH7PS *Sce*DAH7PS and *Nme*DAH7PS (PDB codes: 1KFL, 1OF6 and 4HSO), indicate that the binding site itself has a very strong hydrophobic character, enabling binding of aromatic amino acids. The residue Ser213 has been shown to aid in selectivity of this enzyme towards feedback inhibition by Phe. Upon mutation to glycine, the enzyme becomes sensitive to inhibition by Tyr and no longer by Phe.[26] Tyr was shown to be able to bind to this Ser213Gly variant by ITC

and approximates the K_D seen for Phe binding to the second site in the *Nme*DAH7PS^{WT} (21.1 ± 0.4 µM vs 22 ± 2 µM).[26]

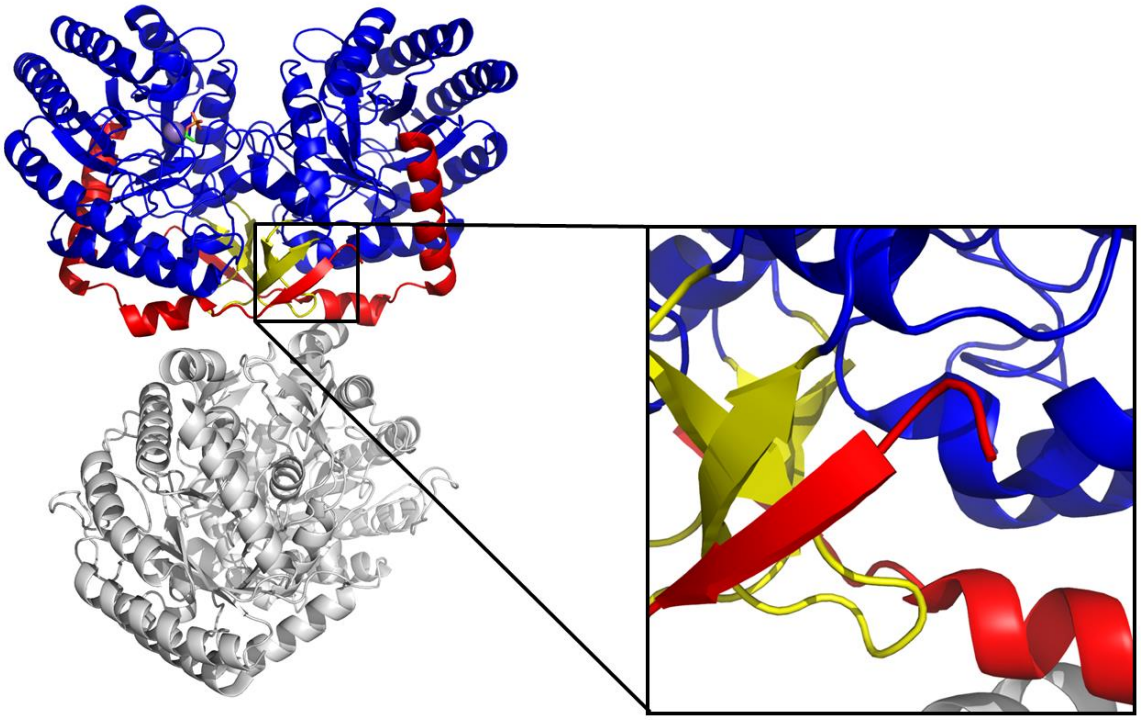


Figure 1.5.4.1 Allosteric binding site in NmeDAH7PS (PDB code 4HSN), illustrating the contribution of the N-terminal extension (red), insertion (yellow) and core barrel (blue)

Comparison of the Tyr-bound *Nme*DAH7PS^{S213G} and the Phe-bound DAH7PS from *E.coli* illustrates some important features of the binding site. Firstly, there is absolute conservation of all residues that interact with the bound ligand. Secondly, the amino acid moiety of the allosteric ligands as well as the residues with which they form hydrogen bonds adopt the same conformation, however, the position of Phe209/211 on the $\beta6_a$ strand is different (PDB codes: 1KFL and 4HSO). In *Eco*DAH7PS, this residue is shifted and forms an end-on π-π stacking interaction with the bound Phe. This interaction is not observed in the Tyr-bound *Nme*DAH7PS^{S213G} variant, which may explain the reduced sensitivity to Tyr observed for this variant compared to the sensitivity of *Nme*DAH7PS^{WT} to Phe (Figure 1.5.4.2). The loss of this interaction may be due to the increased conformational flexibility allowed by Gly in the *Nme*DAH7PS^{S213G} variant.

The Phe-bound *Eco*DAH7PS structure showed that upon Phe binding, subtle conformational changes are propagated across the tight dimer. These changes extend all the way to the catalytic $\beta_2\alpha_2$ loop, which is responsible for the binding of both substrates.[28] In this structure, PEP adopted a flipped conformation. An important structural change noted was the loss of interaction between the $\beta_3\alpha_3$ and $\beta_2\alpha_2$ loops in this structure, hypothesised to contribute to the partial disorder of the loop. Interestingly, the Tyr-bound structure from *Sce*DAH7PS showed similar features, however in this structure, PEP was absent.[20] Both of these structures illustrate the importance of the interaction between the β-hairpin insertion, which makes up a large portion of the allosteric binding site and the catalytic $\beta_2\alpha_2$ loop. The lack of large conformational change observed in the structurally characterised enzymes from the type Iα sub-family in the presence of allosteric ligand implies that subtle variations at the allosteric site may have large effects on possible communication pathways that link the remote allosteric and active sites.

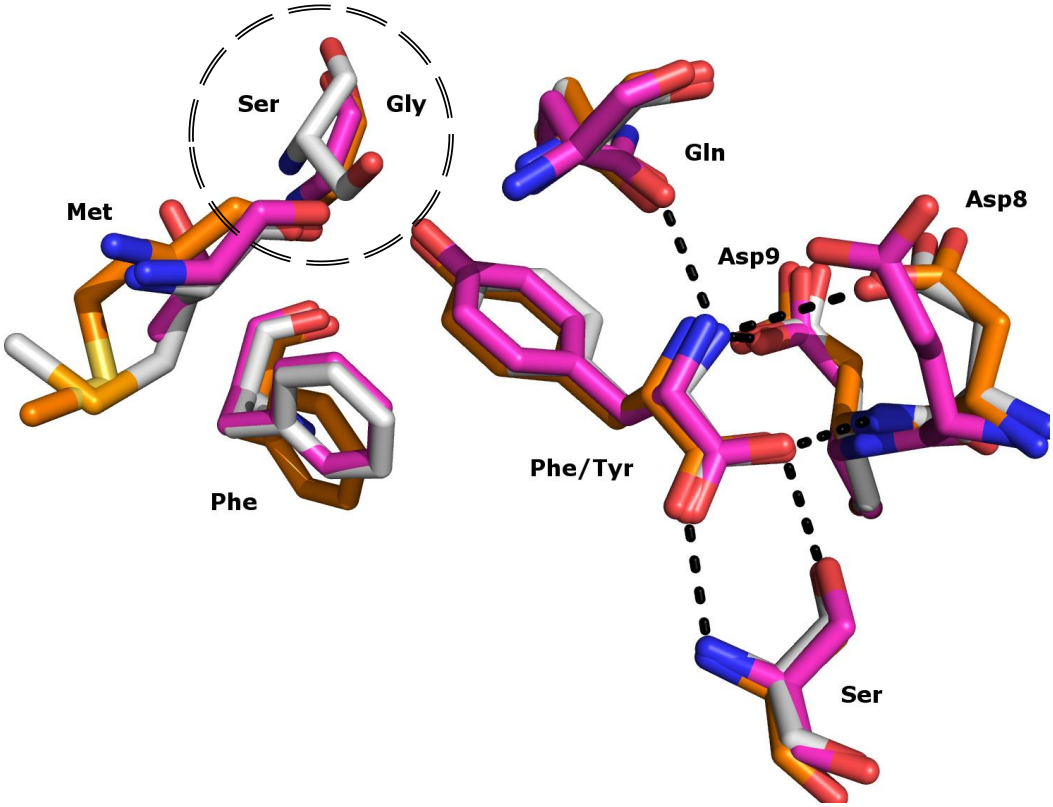


Figure 1.5.4.2 Overlay of the allosteric binding sites of the Tyr bound NmeDAH7PS S213G variant (orange, PDB code 4HSO), the Phe-bound EcoDAH7PS (white, PDB code 1KFL)and the Tyr bound SceDAH7PS(PDB code 1OF6) illustrating hydrogen bonds (black dashed lines) and the single amino acid change that changes selectivity from Phe to Tyr.

1.6 AIMS OF THIS THESIS

The primary aim of this thesis is to better understand how *Nme*DAH7PS is inhibited and regulated. An understanding of inhibition at the active site as well as the allosteric site can aid in future antibiotic development. These studies will disentangle how this enzyme achieves regulation through a dynamic allosteric mechanism involving minimal conformational change.

The specific research objectives of this thesis are:

- To gain a critical understanding of the structure-activity relationships (SAR) between small molecule active site inhibitors of *Nme*DAH7PS
- To understand the reasons why *Nme*DAH7PS is primarily inhibited by Phe.
- To understand how substrates bind at the active site in the presence of allosteric effector.
- To explore the necessity for interactions between the tight dimers at the tetramer interface in *Nme*DAH7PS.
- To identify the roles of hydrophobic residues in the allosteric site and their relative contributions to binding of Phe.
- To gain an in depth understanding for how allostery is communicated between allosteric and active sites of *Nme*DAH7PS without conformational change, accessory domains or complex formation.
- To identify the relative contributions of ionisable residues believed to be responsible for communication of the allosteric signal.

The Phe-sensitive DAH7PS from *E.coli* is the most extensively characterised type Iα DAH7PS to date and shares high sequence identity with *Nme*DAH7PS. For this reason, comparisons between the DAH7PS from these organisms will be referenced throughout this thesis. Comparisons to the Tyr sensitive DAH7PS from *S. cerevisiae,* which has also been well characterised will also be made, to give context to the aforementioned enzymes. This thesis

entails a comprehensive structural and functional characterisation of *Nme*DAH7PS with a view to be able to apply the findings and methods to other systems in nature.

CHAPTER 2- STRUCTURAL ANALYSIS OF SUBSTRATE-MIMICKING INHIBITORS IN COMPLEX WITH *NEISSERIA MENINGITIDIS* 3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE - THE IMPORTANCE OF ACCOMMODATING THE ACTIVE SITE WATER

2.1 INTRODUCTION

Inhibitors of DAH7PS have previously been focussed on the mimicry of substrates or intermediates of the DAH7PS catalysed reaction. Recently published work has involved the design and synthesis of small three-carbon PEP/intermediate mimics of the planarity and stereochemistry of the postulated oxocarbenium and tetrahedral phosphohemiketal intermediate.[53] This work has been developed into the design and synthesis of tetrahedral intermediate mimics with 6-8 carbon atoms, which were shown to inhibit *Mtu*DAH7PS in the nano-molar range.[54]

The previous work on the type Iα DAH7PS from *E. coli* showed that the three-carbon mimics showed variable potency, dictated by their stereochemical makeup. It was found that the planar *E*-vinyl phosphonate was the most potent, followed by the *R*-phospholactate and then the *S*-phospholactate, which exhibited the weakest potency.[51] However, there was no structural investigation into the binding modes of these inhibitors and therefore, an understanding of their relative potencies was theoretical and speculative. The less labile character of the *E*-vinyl phosphonate, which was the most potent inhibitor of *E.coli,* may be highly beneficial as the instability of the phosphate moiety, present in all other designed inhibitors may be unfavourable due to the presence of phosphatases in most biological fluids.[94] Therefore, inhibitors containing such labile functionalities are likely to be problematic during administration of the drug.

This study involved the kinetic and structural characterisation of these three substrate mimics against a type Iα DAH7PS from a pathogenic *Neisseria meningitidis* bacterium in order to provide some insight as to why a planar mimic may have greater affinity for the site. A critical examination of their binding modes and important interactions within the active site will hopefully aid in further modification and extension into inhibitors with greater potency.

This chapter comprises a single publication detailing the inhibition of *N. meningitidis* DAH7PS by three different, three-carbon substrate and intermediate mimics:

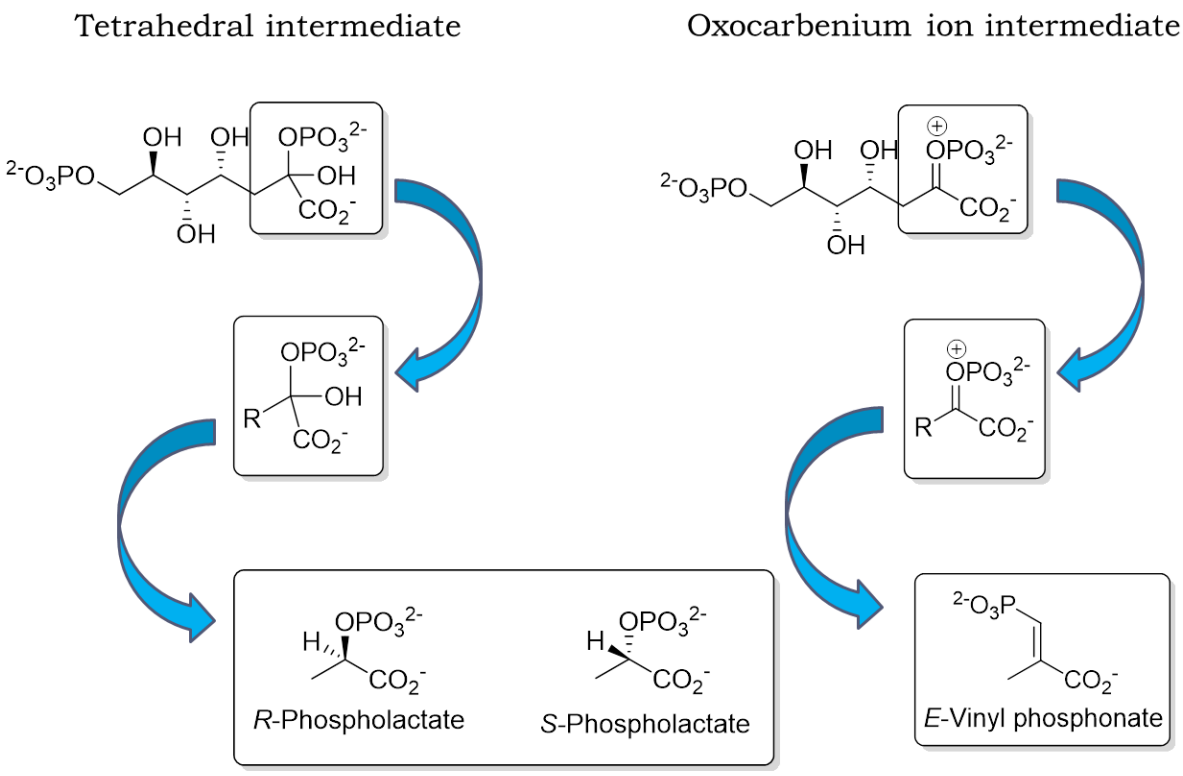


Figure 2.1.1. Rationale for inhibitors to be tested. Phospholactates based on tetrahedral intermediate and vinyl phosphonate based in planar oxocarbenium ion intermediate.

The studies were published in *Bioorganic Chemistry*:

Heyes, L. C., Reichau, S., Cross, P. J., Jameson, G. B. & Parker, E. J. Structural analysis of substrate-mimicking inhibitors in complex with Neisseria meningitidis 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase - The importance of accommodating the active site water. *Bioorg. Chem.* **57,** 242–50 (2014).

Dr S. Reichau aided in the synthesis of the three-carbon compounds. Dr P. J. Cross cloned and expressed *Nme*DAH7PS. Prof G. B. Jameson contributed to the crystallographic analysis of the three co-crystallised inhibitors. Prof E.J. Parker designed experiments and contributed to the writing of the manuscript. The author of this thesis wrote the publication manuscript and all co-authors contributed to the writing and/or revision of the manuscript. The author of this thesis carried out the synthesis, characterisation and testing of the three inhibitors and the growth of crystals, refinement and deposition of the crystal structures reported in the manuscript was also carried out by the author of this thesis.

Bioorganic Chemistry 57 (2014) 242–250

Contents lists available at ScienceDirect

Bioorganic Chemistry

journal homepage: www.elsevier.com/locate/bioorg

Structural analysis of substrate-mimicking inhibitors in complex with *Neisseria meningitidis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase – The importance of accommodating the active site water

Logan C. Heyes [a], Sebastian Reichau [a,1], Penelope J. Cross [a], Geoffrey B. Jameson [b], Emily J. Parker [a,*]

[a] Biomolecular Interaction Centre and Department of Chemistry, University of Canterbury, Christchurch, New Zealand
[b] Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand

ARTICLE INFO

Article history:
Available online 27 August 2014

Keywords:
DAH7PS
DAHPS
Aromatic amino acids
Shikimate pathway
Oxocarbenium ion
Enzyme inhibitors
Meningitis

ABSTRACT

3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) catalyses the first committed step of the shikimate pathway, which produces the aromatic amino acids as well as many other aromatic metabolites. DAH7PS catalyses an aldol-like reaction between phosphoenolpyruvate and erythrose 4-phosphate. Three phosphoenolpyruvate mimics, (*R*)-phospholactate, (*S*)-phospholactate and vinyl phosphonate [(*E*)-2-methyl-3-phosphonoacrylate], were found to competitively inhibit DAH7PS from *Neisseria meningitidis*, which is the pathogen responsible for bacterial meningitis. The most potent inhibitor was the vinyl phosphonate with a K_i value of 3.9 ± 0.4 μM. We report for the first time crystal structures of these compounds bound in the active site of a DAH7PS enzyme which reveals that the inhibitors bind to the active site of the enzyme in binding modes that mimic those of the predicted oxocarbenium and tetrahedral intermediates of the enzyme-catalysed reaction. Furthermore, the inhibitors accommodate the binding of a key active site water molecule. Together, these observations provide strong evidence that this active site water participates directly in the DAH7PS reaction, enabling the facial selectivity of the enzyme-catalysed reaction sequence to be delineated.

1. Introduction

3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) catalyses the first committed step of the shikimate pathway, which is responsible for the biosynthesis of the aromatic amino acids and other important aromatic metabolites in plants and microorganisms [1,2]. In mammals these essential metabolites are obtained from dietary sources, and the enzymes of the shikimate pathway are absent. Therefore, the enzymes of this pathway have been identified as promising targets for the development of new antibacterial therapeutics [3–5].

DAH7PS catalyses the condensation between D-erythrose 4-phosphate (E4P, **2**) and phosphoenol-pyruvate (PEP, **1**) to generate DAH7P (**3**, Fig. 1). Many of the key mechanistic details of the reaction have been elucidated through various labelling, structural and alternative-substrate studies [3,6–12]. The ordered sequential reaction proceeds stereospecifically with respect to both substrates, with the *si* face of PEP attacking the *re* face of E4P. The aldol-like reaction takes place with cleavage of the C–O bond of the phosphate group of PEP, requiring water to act as a co-substrate. The initial nucleophilic attack of PEP on E4P is predicted to be promoted by coordination of the E4P aldehydic oxygen to the essential divalent metal ion, giving an oxocarbenium ion intermediate **4** (Fig. 1) [3,6]. The addition of an active site nucleophilic water results in a phosphohemiketal tetrahedral intermediate **5**, which forms acyclic DAH7P on the elimination of phosphate. The stereochemistry of the tetrahedral intermediate **5**, which is lost on the elimination of phosphate, is determined by the face of the oxocarbenium ion **4** that is attacked by the nucleophilic water. Therefore, the ability of the active site to accommodate and

Abbreviations: BTP, 1,3-bis(tris(hydroxymethyl)methylamino)propane; DAH7PS, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase; E4P, D-erythrose 4-phosphate; EDTA, ethylenediaminetetraacetic acid; LE, ligand efficiency; NCS, non-crystallographic symmetry; NMR, nuclear magnetic resonance; PEP, phosphoenolpyruvate; RMSD, root-mean-square deviation; TMAO, trimethyl-amino-N-oxide; Tris, tris(hydroxymethyl)aminomethane.

* Corresponding author. Address: Department of Chemistry, University of Canterbury, Private Bag 4800, Christchurch, New Zealand. Fax: +64 3 364 2110.
E-mail address: emily.parker@canterbury.ac.nz (E.J. Parker).
[1] Current address: Technische Universität Berlin, Fakultät II Mathematik und Naturwissenschaften, Institut für Chemie / OC / Biologische Chemie, Müller-Breslau-Straße 10, L3, 10623 Berlin, Germany.

http://dx.doi.org/10.1016/j.bioorg.2014.08.003
0045-2068/

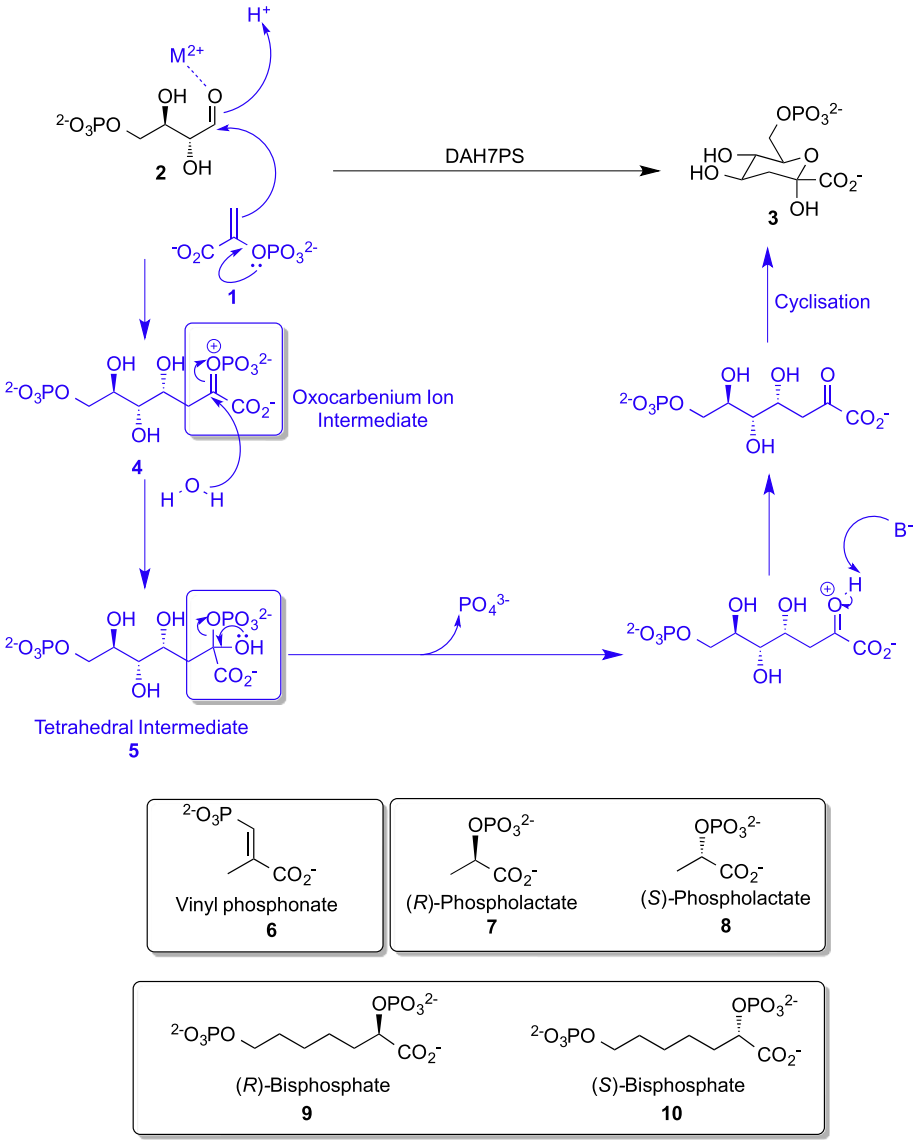


Fig. 1. The overall DAH7PS catalysed reaction showing the likely mechanism via oxocarbenium ion and tetrahedral intermediates (in blue). Compounds **6**–**10** have been shown to act as inhibitors of the DAH7PS enzymes from *Escherichia coli* or *Mycobacterium tuberculosis*.

activate a water molecule is an important aspect of DAH7PS catalysis and, furthermore, may have implications for inhibitor design [3,13].

DAH7PS enzymes have been structurally and functionally characterised from a variety of sources [6,13–19]. While all known enzymes have similar active site architectures supporting a common catalytic mechanism and share a common $(\beta/\alpha)_8$-barrel core, there is considerable sequence divergence between enzymes from different sources leading to the classification of DAH7PS enzymes into two types [20]. Type I enzymes are further divided into type Iα and type Iβ subfamilies, which differ in sequence and their allosteric regulatory machinery. The only structurally characterised type II enzyme is from *Mycobacterium tuberculosis* [16,21].

Active site inhibitors have been previously reported for the phenylalanine-sensitive type Iα DAH7PS from *Escherichia coli* (*Eco*DAH7PS) and the type II enzyme from *M. tuberculosis* (*Mtu*DAH7PS) [3,12,22–24]. PEP-like analogues were found to inhibit *Eco*DAH7PS, with the most potent inhibitor being the vinyl phosphonate, (*E*)-2-methyl-3-phosphonoacrylate **6**, followed by the (*R*)-phospholactate **7** [23]. (*S*)-Phospholactate **8** was the poorest inhibitor. The rationale for a ten-fold difference in potency of the two phospholactate stereoisomers is not clear, as no structures of the inhibited enzymes were determined, but may relate to the stereochemistry of the transient phosphohemiketal tetrahedral intermediate **5**. The vinyl phosphonate **6** was found to have a K_i of 4.7 ± 0.7 μM, which approaches the K_m for PEP of 2 μM for this enzyme [23]. Inhibition studies on the type II *Mtu*DAH7PS were carried out using extended (*R*)- and (*S*)-phospholactates of variable chain length bearing an additional phosphate moiety in order to better mimic the key functionality of the tetrahedral intermediate (bisphosphates **9** and **10**) [3,24]. In each case the compounds bearing (*R*)-configuration were found to inhibit enzymatic activity more potently than the respective (*S*)-enantiomers. The seven-carbon containing (*R*)-bisphosphate **9** had a K_i value of 0.36 ± 0.05 μM, consistent with this compound significantly out-competing the natural substrate PEP (K_m^{PEP} of 37 μM) [3,24].

The genome of *Neisseria meningitidis*, the causative agent of pyogenic meningitis and meningococcal septicaemia, encodes a single type Iα DAH7PS (*Nme*DAH7PS) [17,25]. The crystal structure of *Nme*DAH7PS, which was recently determined in our laboratory, revealed that the active site of *Nme*DAH7PS shares many similarities to the DAH7PS from *E. coli* [33]. The active site Mn^{2+}

is coordinated in a trigonal–bipyramidal fashion by four conserved protein residues, Cys63, His270, Asp304 and Asp324 [17], leaving one coordination site available for the aldehyde moiety of E4P. PEP binding occurs through interaction with the highly conserved residues Arg167, Arg236, Lys188, Arg94 and Lys99. The predicted E4P binding site consists of a highly conserved motif involving the residues Lys99, Pro100, Arg101 and Thr102.

Here we examine the inhibition of *Nme*DAH7PS by the active site directed PEP-mimicking inhibitors **6**, **7** and **8** in order to assess their efficacies against a DAH7PS enzyme from an important pathogen. Furthermore, we report for the first time crystal structures of a DAH7PS, here *Nme*DAH7PS, in complex with each of these inhibitors, revealing that these inhibitors do accurately mimic the binding modes predicted for the oxocarbenium ion **4** and tetrahedral intermediates **5**, along with the co-coordination of a water molecule. Our findings allow refinement of the DAH7PS mechanism, aid the identification of the nucleophilic water in this type Iα enzyme and guide further structure-based inhibitor design of this enzyme.

2. Experimental

2.1. Inhibitor synthesis

(*R*)-Phospholactate **7**, (*S*)-phospholactate **8** and the vinyl phosphonate **6** were synthesised according to the previously reported methods [23]. Concentration determination of the vinyl phosphonate **6** was achieved by comparison of signal integrals from ^{31}P NMR spectroscopy using 85% phosphoric acid as a standard at sufficiently long relaxation times. Phospholactate concentrations were determined via Lanzetta assay for phosphate detection as previously described [3,26].

2.2. Inhibition assays

Enzyme activity was measured by monitoring the loss of absorbance at 232 nm indicating the consumption of PEP (Sigma–Aldrich). Assays were carried out in 50 mM bis(tris(hydroxymethyl) methylamino)propane (BTP) buffer at pH 6.8 with 10 μM ethylenediaminetetraacetic acid (EDTA). The buffer was made up in ultrapure water that had been treated with Chelex resin. E4P was synthesised as reported previously from glucose 6-phosphate [27]. Accurate substrate concentrations were determined by the measurement of absorbance changes that corresponded to the total consumption of substrates in triplicate. The reaction components BTP buffer, $MnSO_4$, PEP, E4P and inhibitor were incubated for two minutes at 25 °C to ensure equilibration before the addition of 2 μL of 0.9 mg mL^{-1} *Nme*DAH7PS. Final assay conditions were 100 μM $MnSO_4$, 115 μM E4P, 19.9–79.6 μM PEP, and variable concentrations of the inhibitors (18.9–151.2 μM vinyl phosphonate **6** inhibitor, 202–1616 μM (*R*)-phospholactate **7** and 268–2144 μM (*S*)-phospholactate **8**). BTP buffer was added to each cuvette to bring the final volume to 1 mL. The values of K_i and K_m were determined by fitting all the initial-rate data points using non-linear regression to the Michaelis–Menten equation modified for the presence of a competitive inhibitor (Eq. (1) below) using Grafit (Erithicus software).

$$v = \frac{V_{max}[PEP]}{K_m(1 + [I]/K_i) + [PEP]} \quad (1)$$

2.3. Co-crystallisation of inhibitors with NmeDAH7PS

*Nme*DAH7PS was overexpressed and purified as previously described [17]. Co-crystallisation of the protein and inhibitor was achieved by preparing a solution containing approximately 5 mM inhibitor dissolved in 10 mM BTP (pH 7.3). 400 μL of this solution was added to the stock solution containing 50 μL of 10 mg mL^{-1} enzyme and was concentrated to 50 μL using a 10 kDa molecular weight cut-off membrane (Vivaspin). The resulting solution was diluted again with 400 μL of inhibitor solution and concentrated by centrifugation. This process was repeated four times. The resulting protein solution was concentrated to approximately 10 mg mL^{-1}. 1 μL of the buffer-exchanged enzyme solution (9–11 mg mL^{-1}) was mixed with 1 μL of crystallisation buffer containing 0.1 M Tris HCl (pH 7.3), 0.2 M trimethyl-amino-*N*-oxide (TMAO), 0.4 mM $MnSO_4$ and 15–20% (w/v) PEG 2000MME. Crystals were grown by hanging-drop vapour diffusion over 500 μL of crystallisation buffer and the crystallisation trays were incubated at 20 °C. Crystals began to form in 24 h and were fully formed within three days. Crystals were flash frozen using liquid nitrogen in a cryoprotectant solution containing reservoir solution and 20% (v/v) PEG400.

2.4. Crystallography and structure determination

X-ray diffraction datasets were collected at the Australian synchrotron using the MX1 and MX2 beamlines [28]. The datasets were integrated and processed using XDS and aimless (CCP4 program suite) [29,30]. The appropriate cut-off resolution was determined via $CC^{1/2} \geqslant 0.5$ while ensuring the data set was complete in the highest resolution shell [31]. For the co-crystal structures with the vinyl phosphonate and (*S*)-phospholactate ligands, the space group ($P12_11$) and unit cell parameters were the same as those previously identified for *Nme*DAH7PS (PDB code 4HSN [17]). To minimise bias, all ligands and waters were removed from the search model (PDB code 4HSN) before molecular replacement (Phaser MR (CCP4) [32]) was carried out. When co-crystallised with (*R*)-phospholactate, *Nme*DAH7PS crystallised, non-isomorphously, in the orthorhombic space group, $P2_12_12_1$. Although structure solution and refinement proceeded normally, the data collected were affected by anisotropy. Refmac5 was used to generate the electron density maps for all three structures and these were manually analysed to develop the structural model by means of COOT [33–35]. The quality of the model was optimised by consecutive model building in COOT and refinement with Refmac5. Water molecules were added manually via interpretation of the $(2|Fo| - |Fc|)$ and $(|Fo| - |Fc|)$ maps, with the requirement that putative water molecules had at least one hydrogen-bonding partner. In all structures, no electron density for the 14 amino acids at the N-terminus of each *Nme*DAH7PS chain could be found, which is attributed to these residues being part of a conformationally highly flexible allosteric binding loop [17]. Residues Met1-Asp9 were also absent from the model (PDB code 4HSN) used for molecular replacement. Molprobity [36] and the validation tools of COOT were used to assess structure quality during refinement cycles and before deposition [36]. After refinement of all protein and non-active-site water molecules had been completed, electron density that could be interpreted as the respective inhibitors could be observed in the active site. CIF files for all three inhibitors were built using the Dundee PRODRG2 server [37]. The inhibitors were positioned manually into this electron density using COOT and then refined. In all three structures, model elaboration and refinement continued until no further improvement in R_{free} could be gained and for final refinements NCS restraints were dropped.

The (*R*)-phospholactate **7** structure crystallised in the orthorhombic space group $P2_12_12_1$ as confirmed by the presence of systematic absences for this space group, and the subsequent successful structure solution and refinement. The data are noticeably anisotropic such that diffraction in the b^* direction is relatively very much weaker than in the other directions, especially at higher

resolution; this is the cause of the high overall R_{merge} (0.298). The low Wilson B value of 19.65 Å^2 vis-à-vis resolution of this dataset is biased by strong diffraction in the other two directions; it is not the result of overlooked twinning, for which no evidence could be found from intensity statistics nor from metrical relationships among unit cell parameters. Notwithstanding anisotropy in the diffraction data, electron-density maps were of noticeably higher quality upon inclusion of the high-resolution data. In this (and the other two structures) the refined average B value was similar to the respective Wilson B value. The final value for R_{free} (0.227) and the small difference between R_{free} and R_{work} (0.026) indicate that this structure (and also the other two structures, which had similarly favourable values) had not been over-refined.

Distances given in the structural analysis of inhibitor binding are defined as a range in Å across all subunits of the tetramer, to ensure consistency during comparison. Only distance differences that remain consistent throughout all subunits during comparison of respective structures are included.

3. Results and discussion

3.1. NmeDAH7PS is inhibited by PEP analogues

Previous studies have shown that both *Eco*DAH7PS and *Mtu*DAH7PS were inhibited by compounds that possess either a planar trigonal or a tetrahedral geometry at C2 [3,22,23]. Both phospholactate stereoisomers **7** and **8** and the vinyl phosphonate **6** were synthesised and tested as inhibitors of *Nme*DAH7PS. All three compounds were found to inhibit *Nme*DAH7PS competitively with respect to substrate PEP. The most potent inhibitor was the vinyl phosphonate **6**, with an observed K_i value of 3.9 ± 0.4 μM. This value is significantly lower than the measured K_m for PEP (12 μM). While both enantiomers of the phospholactate were also inhibitory, inhibition by the (*R*)-phospholactate **7** was 3.5 times greater than that observed for the (*S*)-phospholactate **8** (Table 2).

The order of inhibitory potency observed for these compounds with *Nme*DAH7PS parallels the observations made for the type Iα *Eco*DAH7PS. For *Eco*DAH7PS, the (*R*)-phospholactate **7** was shown to be more than 10-fold more potent than the (*S*)-stereoisomer [23]. The *Mtu*DAH7PS enzyme also demonstrated this same stereochemical preference for the configuration at C2 when tested with the extended bisphosphate inhibitors **9** and **10** [3,24]. The most potent of these bisphosphate inhibitors was the (*R*)-bisphosphate **9**, which exhibited a K_i value of 0.36 ± 0.05 μM (in comparison to the (*S*)-bisphosphate **10** with a K_i of 0.6 ± 0.1 μM) [3]. Thus, stereochemical influences on inhibitor-binding affinity appear to be intrinsic to the active sites of DAH7PS and are independent of the class of DAH7PS and the species from which the enzyme is derived.

The increase in potency upon introduction of a phosphonate with a planar geometry at C2, observed for vinyl phosphonate **6**, is consistent with what has been reported previously for *Eco*DAH7PS [23]. It should be noted that all these inhibitors are particularly potent with respect to their size, showing very high ligand efficiencies (ligand efficiency (LE) = $\Delta G/N$; where $\Delta G = -RT\ln K_i$ and N = number of non-hydrogen atoms per ligand) [38,39]. When comparing the relative ligand efficiencies of the vinyl phosphonate measured here for *Nme*DAH7PS, and the extended (*R*)-bisphosphate **9** (as an inhibitor of *Mtu*DAH7PS), it is evident that the vinyl phosphonate **6** is 1.6-fold more efficient than the longer inhibitor. The (*R*)-bisphosphate **9** is the most potent inhibitor reported for any DAH7PS enzyme to date [3].

3.2. Structural analysis of inhibitor binding

Structures of *Nme*DAH7PS co-crystallised with each of the inhibitors were determined by X-ray diffraction. The three structures containing the ligands vinyl phosphonate **6**, (*R*)-phospholactate **7** and (*S*)-phospholactate **8** were refined to resolutions of 1.76 Å, 2.17 Å and 2.35 Å respectively (Table 1). The vinyl phosphonate and (*S*)-phospholactate-bound proteins crystallised in the same space group (monoclinic $P12_11$) as the wild-type protein, whereas the (*R*)-phospholactate-bound enzyme crystallised in the orthorhombic space group $P2_12_12_1$. The asymmetric unit for all three inhibitor-bound structures contained a homotetramer.

Continuous electron density for the three ligands was observed in all active sites after complete refinement of the protein and non-active-site water molecules (Fig. 2). Judging from B values of ligands and ligand-binding side chains, the ligands are present in full, or close to full, occupancy. The binding of inhibitors does not affect the secondary or tertiary structure of the enzyme. Individual subunits within a tetramer superimpose on each other with a root-means-square displacement (RMSD) of between 0.05 and 0.09 Å, whereas superposition of subunits from different tetramers superimpose with RMSD values in the range 0.10 to 0.18 Å.

The inhibitors show similar placement within the active site and binding modes to that observed for the natural substrate PEP **1** (Fig. S4, Supplementary material). The side chains of residues Lys188, Arg167 and Arg236 and the main-chain peptide N of Ala166 that interact with the phosphate moiety of PEP form similar interactions with the phosphate or phosphonate functionalities of the three ligands, consistent with this portion of each inhibitor being able to mimic the PEP binding characteristics well. Mn^{2+} is found in the metal-binding site of all the structures in full occupancy. As well as coordinating to the Mn^{2+} in the active site, the carboxylate group of each ligand forms salt bridges with the side chain of Arg94 and with the side chain of Lys99 of the conserved LysProArgThr motif, which extends into the E4P binding site. In the structure of the PEP-bound enzyme, Mn^{2+} also coordinates the PEP carboxylate functionality and a water molecule. The lack of this coordinating water molecule in any of the inhibitor-bound structures is associated with a closer positioning of the carboxylate group of the respective PEP-mimic to the metal ion (Fig. S2).

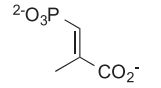
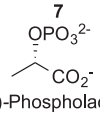
There are five waters in the active site of each subunit occupying similar positions in the vicinity of both the vinyl phosphonate **6** and (*R*)-phospholactate **7** inhibitors (labelled WAT1-5, Fig. 2). WAT1 is found in a position corresponding to the *re* face of PEP and establishes hydrogen bonds with the bridging oxygen of the phosphate group and with a terminal oxygen of the phosphate or phosphonate functionalities. Additionally, WAT1 hydrogen-bonds to WAT2 and Glu145. WAT2 also forms hydrogen-bonding interactions with Lys188 and Tyr96. WAT3 is also hydrogen-bonded to the phosphate or phosphonate moieties and interacts with the backbone amino group and the η^1 nitrogen of Arg167. WAT4 forms hydrogen-bonding interactions with both η^1 and η^2 of Arg167 and to an oxygen of the phosphate or phosphonate functionalities. WAT5 makes no direct interactions with the inhibitors, but provides a hydrogen-bonding bridge between Glu145 and Lys99. WAT1, WAT2, WAT3 and WAT5 are also found occupying similar positions in the (*S*)-phospholactate **8** inhibitor-bound structure. No evidence for a WAT4 was observed for this structure, although it is noted that this structure has the lowest resolution of this set. All five of these water molecules were also present in the PEP-bound structure, although WAT4 is only modelled into chain A of this structure [17]. In addition, as noted earlier, a Mn^{2+}-coordinated water is observed in the PEP-bound structure.

Tertiary structures, as noted above, are unaffected by inhibitor binding and the tetrameric quaternary structures of *Nme*DAH7PS bound with (S)-phospholactate and vinyl phosphonate are identical to that of the isomorphous PEP-bound parent structure (PDB code 4HSN) (RMSDs for superposition onto 4HSN of 0.17 and 0.23 Å for 1328 and 1327 Cα atoms, respectively). However, there are small but significant changes in the arrangement of subunits

Table 1
Data collection and refinement.

	*Nme*DAH7PS complex		
	Vinyl phosphonate **6**	(*R*)-phospholactate **7**	(*S*)-phospholactate **8**
Data Collection			
Crystal system; space group	Monoclinic, $P12_11$	Orthorhombic, $P2_12_12_1$	Monoclinic, $P12_11$
Unit cell parameters			
a, b, c (Å)	73.49, 136.50, 76.16	79.44, 133.7, 147.2	73.42, 137, 76.48
α, β, γ (°)	90, 96.41, 90	90, 90, 90	90, 96.48, 90
Resolution range (Å)	46.82–1.76 (1.79–1.76)	48.31–2.17 (2.21–2.17)	46.88–2.34 (2.40–2.34)
Measurements	1,113,629	617,846	242,429
Unique reflections	147,006	83,610	63,250
Redundancy	7.6	7.4	3.8
Completeness (%)	99.8 (99.2)	99.9 (98.7)	99.9 (99.8)
I/σ (*I*)	16.9 (1.9)	6.5 (1.2)	8.0 (1.1)
R_{merge}	0.070	0.298	0.148
$CC_{1/2}$	0.88	0.52	0.50
Wilson *B* value $(Å)^2$	24.52	19.65	30.63
Matthews coefficient	2.50	2.57	2.51
Refinement			
R_{work}	0.1918	0.2007	0.2040
R_{free}	0.2144	0.2267	0.2304
Chain length	351	351	351
Observed number of residues	333 (Chain C & D) 332 (Chain A & B)	333 (Chain A), 332 (Chain B, C &D)	333 (Chain A & B), 334 (Chain C), 332 (Chain D)
Water molecules	779	634	263
Other	4	8	8
Ligand	4	4	4
Mean B $(Å)^2$			
Protein	32.87	24.90	38.37
Water	34.50	24.42	30.42
Other	23.66	33.12	57.93
Ligand	24.35	22.54	40.23
R.m.s.d from target values			
Bond lengths (Å)	0.0104	0.0106	0.0117
Bond angles	1.3822	1.3951	1.4594
Dihedral angles	0.0800	0.0745	0.0761
Ramachandran			
Preferred (%)	97.71	97.54	98.36
Allowed (%)	1.68	1.84	1.22
Outliers (%)	0.61	0.61	0.61
PDB Entry	4UMA	4UMB	4UMC

Table 2
Inhibition constants of compounds tested against *Nme*DAH7PS.

Inhibitor	K_i (μM)[a]
 Vinyl phosphonate **6**	3.9 ± 0.4
 (*R*)-Phospholactate **7**	99 ± 16
 (*S*)-Phospholactate **8**	360 ± 30

[a] Assay conditions used were 100 μM $MnSO_4$, 115 μM E4P, 19.9–79.6 μM PEP, 18.9–151.2 μM vinyl phosphonate inhibitor **6**, 202–1616 μM (*R*)-phospholactate **7**, and 268–2144 μM (*S*)-phospholactate **8**. A graphical representation of the data is provided in the supplementary material (Fig. S1 and Table S1).

for the non-isomorphous (*R*)-phospholactate structure relative to the PEP-bound parent (RMSD of 0.60 Å for superposition of 1325 Cα atoms). The tetramer is made up of two dimers, one comprising an extensive interface with mutual proximity of the E4P binding loops (β2α2), often referred to as the tight dimer interface, the other dimer, often referred to as the loose dimer interface, being formed largely by antiparallel association of helix α0. The quaternary structural changes are concentrated at the weak dimer interface, the tight dimers superimposing on PEP-bound enzyme with RMSDs in the range 0.13–0.27 Å. Thus, these differences between the (*R*)-phospholactate-bound structure and the PEP-, (*S*)-phospholactate- and vinyl phosphonate-bound structures can be attributed to crystal packing effects (sampling different conformations of the ensemble of quaternary structures freely sampled in solution).

3.3. Small differences in inhibitor binding modes are observed

While the binding modes of the two phospholactate enantiomers are very similar, a comparison of the (*R*)- and (*S*)-phospholactate inhibitor-bound structures and the different way in which the two stereoisomers can be accommodated in the active site highlights some small differences in the relative positions of the key functional groups which may account for their inhibitory differences (Fig. 3). These differences are shared when any subunit of one tetramer is compared to any subunit of another tetramer, consistent with the very high similarity of all individual subunits described above. This allows small differences of marginal significance in a single pairwise comparison to become highly significant when all subunits are compared. The phosphate groups of the two molecules are similarly positioned and a slight displacement of the carboxylate functionality of (*R*)-phospholactate away from Lys99 towards Arg94 is observed. WAT1 is found in closer proximity to

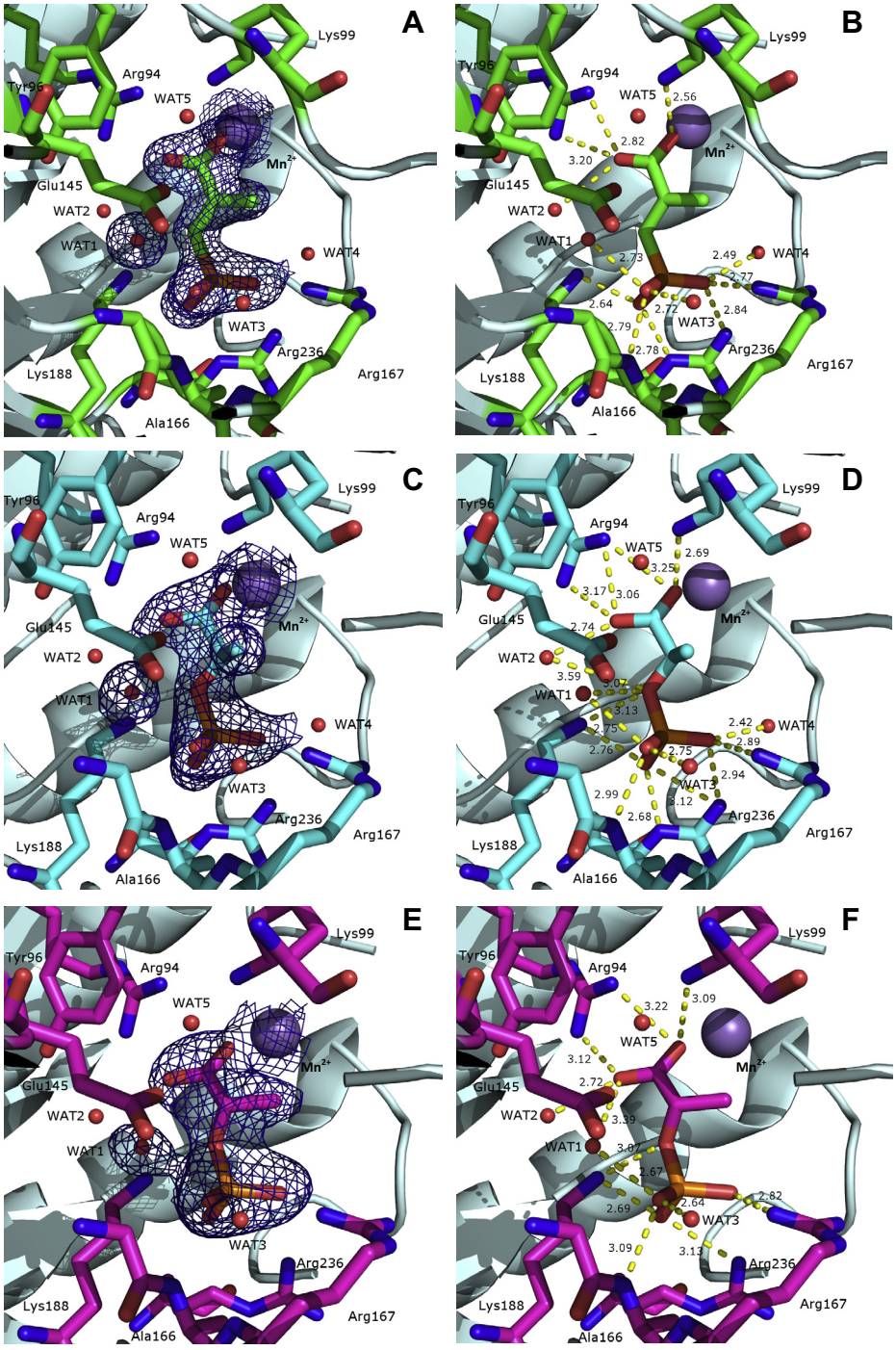


Fig. 2. Structure of *Nme*DAH7PS in complex with inhibitors. Panels A and B: vinyl phosphonate **6** bound structure (chain A). Panels C and D: (*R*)-phospholactate **7** bound structure (chain A). Panels E and F: (*S*)-phospholactate **8** bound structure (chain A). Panels B, D and F show interatomic distances (in Å). Panels A, C and E display the $|2Fo - Fc|$ map to indicate the observed electron density for the bound ligand and WAT1 contoured at 1σ. The interatomic distances for each chain and the omit maps are provided in the supplementary material (Figs. S3–S7). Mn^{2+} is shown as a purple sphere in all panels.

C2 of the (*S*)-phospholactate than the (*R*)-phospholactate (by 0.2–0.7 Å by comparison across all chains), bringing this separation to less than van der Waals distance. This is likely to be a direct result of the configurational difference between the two phospholactate inhibitors. The stereochemical difference will necessarily result in the C2 hydrogen of the inhibitor pointing directly towards WAT1 when (*S*)-phospholactate is bound, whereas for the (*R*)-enantiomer, this hydrogen will be positioned away from WAT1 (Fig. 4). In addition, the positioning of the (*R*)-phospholactate allows for more optimal hydrogen-bonding interactions to WAT1, as can be seen when the angles (C2 phospholactate-bridging phosphate oxygen-WAT1) are compared between the (*R*)- and (*S*)-phospholactate inhibitor-bound structures: for (*S*)-phospholactate, this angle is close to 90°, leading to poor hydrogen-bonding interactions when the direction of the lone pair electrons of the bridging oxygen is considered. In contrast, the angle (C2 phospholactate-bridging phosphate oxygen-WAT1) is close to 109° in the (*R*)-phospholactate bound structure, which leads to a close to optimal angular geometry for a hydrogen bond between WAT1 and the bridging phosphate oxygen in the (*R*)-phospholactate ligand. In

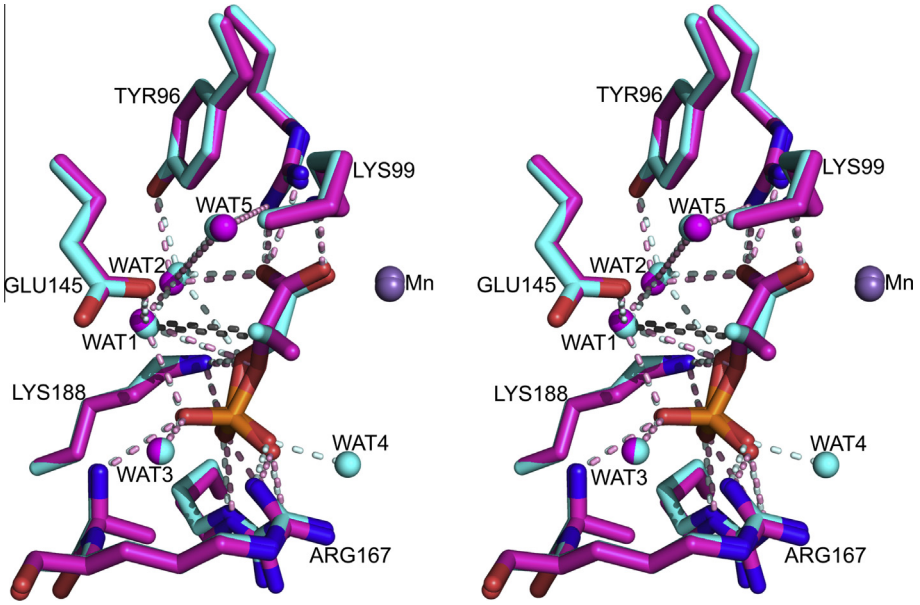


Fig. 3. Comparison of binding modes of (*R*)- and (*S*)-phospholactate in stereo view. (*R*)-Phospholactate is shown in cyan and (*S*)-phospholactate is shown in magenta. Interatomic distances are shown as pale cyan dashes for the (*R*)-phospholactate interactions and pink dashes for the (*S*)-phospholactate interactions. Distances between WAT1 and C2 of the ligand are 3.8 Å for (*R*)-phospholactate and 3.3 Å for the (*S*)-phospholactate (shown as grey dashes). Mn^{2+} is shown as a purple sphere. The carboxylate and phosphate moieties of both inhibitors are in very similar positions, which correspond well to the placement of these moieties when the substrate PEP is bound; the positions of ligand-binding side chains do not change significantly on ligand binding. However, a slight displacement of the bridging phosphate oxygen is apparent, along with the methyl substituent, which results in an altered hydrogen-bonding geometry between WAT1 and the bridging phosphate oxygen.

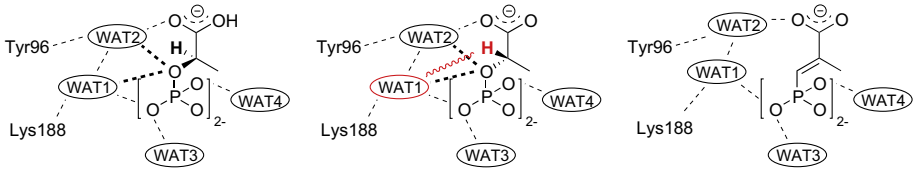


Fig. 4. Schematic representation of subtle differences in the water-mediated hydrogen-bonding network in the inhibitor bound crystal structures. For (*R*)-phospholactate (left), the C2–hydrogen (bold) points away from WAT1. For (*S*)-phospholactate (centre), the C2–hydrogen (bold and red) points towards WAT1, potentially resulting in steric hindrance (indicated by red wavy line). Both phospholactate enantiomers establish hydrogen-bonding interactions with WAT1 and WAT2 through the bridging phosphate oxygen (dotted lines). These interactions are not observed in the binding mode of *E*-vinyl phosphonate (right), which has a bridging methylene group instead of an oxygen.

summary, these small differences in the co-crystal structures of *Nme*DAH7PS in complex with the two enantiomers of phospholactate, improved accommodation of WAT1 and a more optimal hydrogen-bonding network between the inhibitor and the active-site waters may serve to explain the difference in potency of the enantiomeric phospholactates. Our study reveals that the small differences in inhibitory potency correlate with subtly different binding modes of the inhibitors as could have been predicted; however, the precise structural information obtained here on the enzyme-inhibitor co-crystal structures will help inform future inhibitor design.

The increased inhibitor potency observed for the vinyl phosphonate **6** over the (*R*)-phospholactate **7** suggests that the active site is better prepared to accommodate a planar rather than tetrahedral intermediate. This occurs despite the lack of a hydrogen bond acceptor in the vinyl phosphonate, which results in one less hydrogen bond being established between the vinyl phosphonate and WAT1. Although a similar binding mode is adopted by both inhibitors and the distances between the phosphonate or phosphate moieties of these inhibitors and the residues with which they interact are very similar, there are small structural differences that may be relevant to this change in functionality and potency. The bridging oxygen of the phosphate in the (*R*)-phospholactate is 0.1–0.4 Å closer to Lys188 than the bridging carbon of the phosphonate functionality in the vinyl phosphonate. A small relative displacement (0.1–0.3 Å) of the carboxylate functionality of the vinyl phosphonate towards to the amino group of Lys99 is also apparent. This shift in the binding mode of the vinyl phosphonate towards Lys99 compromises the interaction between the carboxylate functionality and WAT2, which are further apart (0.2–0.5 Å) in the vinyl phosphonate bound structure. However, the planar and more compact geometry of the vinyl phosphonate compared to (*R*)-phospholactate enables closer proximity and thus a stronger Coulomb interaction to develop between the negative charge of the vinyl phosphonate carboxylate and the potentially positively charged Lys99 amino group. Another key factor in the stronger binding of the vinyl phosphonate compared to the phospholactate moieties is that the rotational/torsional degrees of freedom around the C1–C2 and the C2–C3 bond are surrendered upon binding to the enzyme for the phospholactate species, in contrast to minimal loss of such degrees of freedom when the more rigid vinyl phosphonate binds. This can be estimated to contribute a free-energy stabilisation as $2RT\ln 3$ (allowing three torsional conformations), that is $\sim$5.4 kJ mol^{-1} at 25 °C, or a factor of $\sim$9 in terms of equilibrium constant, a number remarkably similar to the relative affinities of binding of the vinyl phosphonate compared to the more easily accommodated (*R*)-enantiomer of the phospholactate species.

The accommodation of WAT1 appears to be an important feature of all the inhibitor-*Nme*DAH7PS structures. The distance

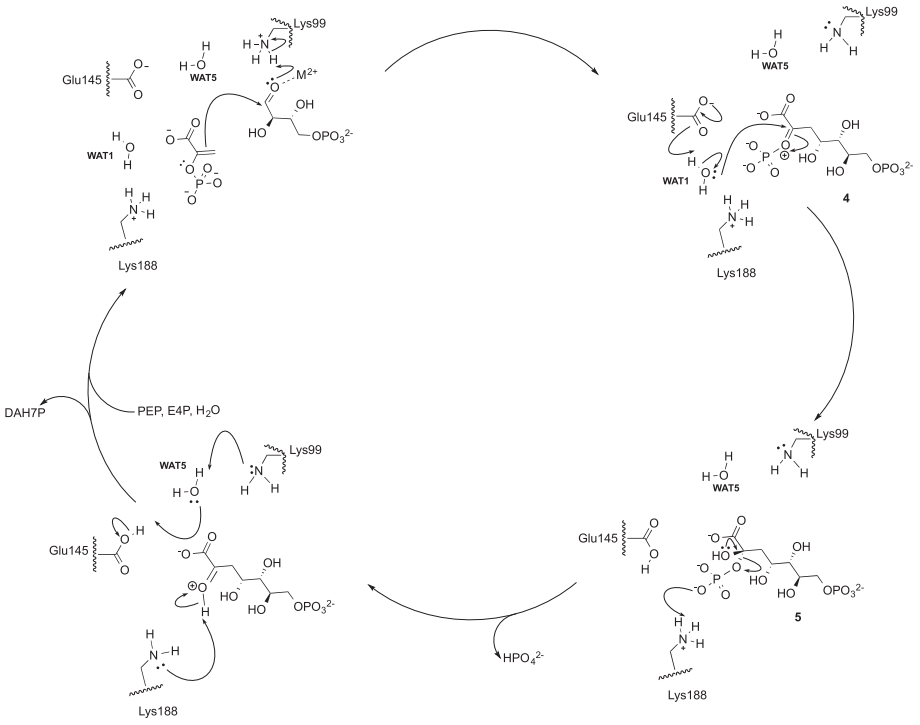


Fig. 5. Proposed mechanism of catalysis of *Nme*DAH7PS showing key nucleophilic role for active site WAT1.

from WAT1 to C2 of the vinyl phosphonate is (0.1–0.3 Å) closer than is that observed for the (*R*)-phospholactate-bound structure. This difference appears to be a direct result of the planarity resulting from the double bond of the vinyl phosphonate concomitant with the fixing of the carboxylate and phosphate moieties into their binding pockets. Taken together with the inhibition constants, this analysis suggests that the geometry of the *Nme*DAH7PS active site is better optimised to accommodate a planar reaction intermediate or mimic (vinyl phosphonate) thereof, in combination with the nucleophilic WAT1. The loss of a hydrogen bonding contact when moving from the phospholactates to vinyl phosphonate appears to be outweighed by WAT1 being able to be accommodated at a slightly better van der Waals separation than the (*R*)-phospholactate, which at 3.8–3.9 Å is slightly longer (and (*S*)-phospholactate which is noticeably shorter) than ideal van der Waals contact, even after allowing for different hybridisation of carbon atom C2.

3.4. Implications for the mechanism of the DAH7PS-catalysed reaction

The cleavage of the C–O bond to release phosphate in the reaction catalysed by DAH7PS demands that water acts as a co-substrate in the reaction (Fig. 1). Coordination of the aldehydic group of E4P to the divalent active site metal ion both assists in presenting the *re* face of the aldehyde group to the *si* face of PEP and thus pre-organises the substrates for nucleophilic attack by PEP through its C3. It is likely that an oxocarbenium-like transition state or transient intermediate is then quenched by water to generate a tetrahedral intermediate. The direction of attack of this water determines the stereochemistry of this intermediate. From the structural information gathered in these ligand-bound structures, it is evident that the most likely candidate for the nucleophilic water is indeed WAT1 based on its proximity to C2 of PEP, supporting the attack by water on the *re* face of PEP, overall *anti* addition across the PEP-double bond, and the generation of a transient stereogenic centre in the intermediate with an (*S*) absolute configuration. In support of this, water has been observed to occupy similar positions in the DAH7PS structures from *Saccharomyces cerevisiae*, *E. coli, M. tuberculosis* and *Thermotoga maritima* [3,6,13,14].

The preference for the binding of the (*R*)-phospholactate over the (*S*)-phospholactate and the structural observations from these ligand-bound structures are also consistent with *re* face water attack and the generation of (*S*)-configuration at C2 of the tetrahedral intermediate. A greater than 3-fold difference in potency between the (*R*)- and (*S*)-phospholactate inhibitors results from change in absolute configuration at C2. Although an (*S*)-configuration of the phosphohemiketal tetrahedral intermediate **5** will arise from *re* face water attack, the closer placement of C2 of the (*S*)-phospholactate and the corresponding requirement to accommodate a hydrogen on C2 of the phospholactate likely hinders co-occupation of the active site with this water molecule, accounting for the reduced potency observed for the (*S*)-phospholactate. Conversely, (*R*)-phospholactate, with the hydrogen at C2 pointing away from the active site water more readily accommodates WAT1. A similar stereochemical preference was also observed for the extended (*R*)- and (*S*)-bisphosphate inhibitors **9** and **10** for *Mtu*DAH7PS. While the preference for the (*R*)-configuration was only twofold, the differences in potency were also attributed to better accommodation of a water molecule in an equivalent position to WAT1 from this study [3].

The planar geometry of the vinyl phosphonate is accompanied by a significant increase in potency over either of the phospholactate stereoisomers. The double bond introduced at C2 is able to mimic the prochirality seen in the oxocarbenium ion intermediate, which further suggests that the planar ligand geometry is best suited to accommodate the binding of the nucleophilic water molecule (WAT1). Therefore, among the three ligands examined in this study, the ability for the inhibitor to adequately accommodate WAT1 upon binding and the lessened changes in conformational flexibility on binding with respect to the free ligand are the defining criteria for inhibitor potency. From our structural analysis, we propose that (*R*)-phospholactate is a stronger inhibitor than (*S*)-phospholactate, in part due to improved hydrogen-bond

geometry between WAT1 and the bridging phosphate oxygen in (*R*)-phospholactate and in part due to lesser stereochemical clash of WAT1 with C2. We further suggest that the highest potency of the planar vinyl phosphonate inhibitor is due in part to lowered loss of conformational flexibility of this more rigid ligand on binding to the enzyme and in part on WAT1 being placed in ideal proximity to C2 of the ligand. This position may most resemble that which WAT1 would assume during the course of regular enzymatic turnover, acting as the nucleophilic water which attacks C2 of the oxocarbenium ion **4** (Fig. 1).

With the identification of the nucleophilic water, the details of the *Nme*DAH7PS-catalysed reaction mechanism can be elucidated from the kinetic and structural data presented here in combination with previous mechanistic and structural studies on DAH7PS from other sources (Fig. 5) [3,10]. Nucleophilic attack of PEP at the E4P aldehyde moiety is facilitated by activation of the carbonyl moiety by coordination to the active-site metal ion and subsequent protonation of the nascent alkoxo group by Lys99. The transient oxocarbenium ion species **4** is then rapidly attacked by WAT1, which is deprotonated by Glu145, resulting in the formation of the tetrahedral intermediate **5**. This nucleophilic water is held in position on the *re* face of PEP by hydrogen bonds to Glu145 and WAT2. Glu145 is conserved in both sequence and structure in all known DAH7PS enzymes [3,13]. Lys99 residue is protonated via an active site water molecule by Glu145. Lys188, which is positioned within proximity for hydrogen bonding interactions with the phosphate of PEP, is predicted to aid in the elimination of phosphate from the tetrahedral intermediate **5** via protonation of the leaving group. Elimination of phosphate results in the linear DAH7P product, which is protonated at the 2-keto moiety. Deprotonation of this species by Lys188 subsequently results in the formation of the product, DAH7P, which is in equilibrium with its pyranose form **3**, and restores the enzyme for a subsequent catalytic cycle (Fig. 5).

4. Conclusion

Three PEP mimics were found to be inhibitors of *Nme*DAH7PS and to occupy the PEP binding pocket of the active site of this enzyme. Analysis of the first co-crystal structures of these small but highly atom-efficient inhibitors with their target enzyme highlight the importance of the accommodation of the nucleophilic water molecule in the active site for potent inhibition. From our studies we conclude that planar rather than tetrahedral PEP-mimics are more promising candidates for inhibition of *Nme*DAH7PS, and the ability of inhibitors to accommodate the key nucleophilic water molecule needs to be considered in future inhibitor design. Extension of a planar ligand based on the rigid vinyl phosphonate scaffold facilitated by the first structural information on enzyme-inhibitor complexes presented here would likely be a highly potent inhibitor of *Nme*DAH7PS.

Acknowledgments

This research was undertaken on the MX1 and MX2 beamlines at the Australian Synchrotron, Victoria, Australia. Funding is gratefully acknowledged for these studies from the New Zealand Marsden Fund (UOC1105). S.R. is grateful for financial support from a University of Canterbury Doctoral Scholarship and a New Zealand International Doctoral Research Scholarship.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bioorg.2014.08.003.

References

[1] R. Bentley, Crit. Rev. Biochem. Mol. Biol. 25 (1990) 307–384.
[2] H. Maeda, N. Dudareva, Annu. Rev. Plant Biol. 63 (2012) 73–105.
[3] S. Reichau, W. Jiao, S.R. Walker, R.D. Hutton, E.N. Baker, E.J. Parker, J. Biol. Chem. 286 (2011) 16197–16207.
[4] V.F.V. Prazeres, L. Tizon, J.M. Otero, P. Guardado-Calvo, A.L. Llamas-Saiz, M.J. van Raaij, L. Castedo, H. Lamb, A.R. Hawkins, C. Gonzalez-Bello, J. Med. Chem. 53 (2010) 191–200.
[5] J.R. Coggins, C. Abell, L.B. Evans, M. Frederickson, D.A. Robinson, A.W. Roszak, A.P. Lapthorn, Biochem. Soc. Trans. 31 (2003) 548–552.
[6] I.A. Shumilin, R. Bauerle, J. Wu, R.W. Woodard, R.H. Kretsinger, J. Mol. Biol. 341 (2004) 455–466.
[7] H.G. Floss, D.K. Onderka, M. Carroll, J. Biol. Chem. 247 (1972) 736–744.
[8] A.B. DeLeo, D.B. Sprinson, Biochem. Biophys. Res. Commun. 32 (1968) 873–877.
[9] D.K. Onderka, H.G. Floss, J. Am. Chem. Soc. 91 (1969) 5894–5896.
[10] M. Ahn, A.L. Pietersma, L.R. Schofield, E.J. Parker, Org. Biomol. Chem. 3 (2005) 4046–4049.
[11] R.M. Williamson, A.L. Pietersma, G.B. Jameson, E.J. Parker, Bioorg. Med. Chem. Lett. 15 (2005) 2339–2342.
[12] S.R. Walker, E.J. Parker, Bioorg. Med. Chem. Lett. 16 (2006) 2951–2954.
[13] V. König, A. Pfeil, G.H. Braus, T.R. Schneider, J. Mol. Biol. 337 (2004) 675–690.
[14] I.A. Shumilin, R.H. Kretsinger, R.H. Bauerle, Structure (London) 7 (1999) 865–875.
[15] L.R. Schofield, B.F. Anderson, M.L. Patchett, G.E. Norris, G.B. Jameson, E.J. Parker, Biochemistry 44 (2005) 11950–11962.
[16] C.J. Webby, H.M. Baker, J.S. Lott, E.N. Baker, E.J. Parker, J. Mol. Biol. 354 (2005) 927–939.
[17] P.J. Cross, A.L. Pietersma, T.M. Allison, F.C. Cochrane, E.J. Parker, Protein Sci. 22 (2013) 1087–1099.
[18] L. Zhou, J. Wu, V. Janakiraman, I.A. Shumilin, R. Bauerle, R.H. Kretsinger, R.W. Woodard, Bioorg. Chem. 40 (2012) 79–86.
[19] S.H. Light, A.S. Halavaty, G. Minasov, L. Shuvalova, W.F. Anderson, Protein Sci. 21 (2012) 887–895.
[20] R.A. Jensen, G. Xie, D.H. Calhoun, C.A. Bonner, J. Mol. Evol. 54 (2002) 416–423.
[21] C.J. Webby, W. Jiao, R.D. Hutton, N.J. Blackmore, H.M. Baker, E.N. Baker, G.B. Jameson, E.J. Parker, J. Biol. Chem. 285 (2010) 30567–30576.
[22] S.R. Walker, W. Jiao, E.J. Parker, Bioorg. Med. Chem. Lett. 21 (2011) 5092–5097.
[23] S.R. Walker, H. Cumming, E.J. Parker, Org. Biomol. Chem. 7 (2009) 3031–3035.
[24] S. Reichau, E.J. Parker, RSC Adv. 3 (2013) 3209–3212.
[25] N.E. Rosenstein, B.A. Perkins, D.S. Stephens, T. Popovic, J.M. Hughes, New England J. Med. 344 (2001) 1378–1388.
[26] P.A. Lanzetta, L.J. Alvarez, P.S. Reinach, O.A. Candia, Anal. Biochem. 100 (1979) 95–97.
[27] C.E. Ballou, H.O.L. Fischer, D. MacDonald, J. Am. Chem. Soc. 77 (1955) 5967–5970.
[28] T.M. McPhillips, S.E. McPhillips, H.J. Chiu, A.E. Cohen, A.M. Deacon, P.J. Ellis, E. Garman, A. Gonzalez, N.K. Sauter, R.P. Phizackerley, S.M. Soltis, P. Kuhn, J. Synchrotron Rad. 9 (2002) 401–406.
[29] P.R. Evans, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 282–292.
[30] P. Evans, Acta Crystallogr. D Biol. Crystallogr. D62 (2006) 72–82.
[31] P.A. Karplus, K. Diederichs, Science 336 (2012) 1030–1033.
[32] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, J. Appl. Crystallogr. 40 (2007) 658–674.
[33] G.N. Murshudov, P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 355–367.
[34] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Acta Crystallogr. Sect. D: Biol. Cryst. 53 (1997) 240–255.
[35] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501.
[36] V.B. Chen, W.B. Arendall III, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, Acta Crystallogr. D Biol. Crystallogr. D66 (2010) 12–21.
[37] A.W. Schuttelkopf, D.M.F. van Aalten, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 1355–1363.
[38] I.D. Kuntz, K. Chen, K.A. Sharp, P.A. Kollman, Proc. Natl. Acad. Sci. USA 96 (1999) 9997–10002.
[39] S.D. Bembenek, B.A. Tounge, C.H. Reynolds, Drug Discovery Today 14 (2009) 278–283.

CHAPTER 3- INVESTIGATIONS INTO THE SPECIFICITY OF THE ALLOSTERIC BINDING SITE AND THE STRUCTURAL REORGANISATION OF THE ACTIVE SITE

3.1 INTRODUCTION

N. meningitidis expresses a single DAH7PS that is primarily inhibited by phenylalanine (Phe).[26] However, two other aromatic amino acids, tyrosine (Tyr) and tryptophan (Trp) also inhibit the enzyme to a lesser extent.[26] This promiscuity at a single binding site of *Nme*DAH7PS has interesting implications for drug design and discovery. Traditionally, drug discovery attempts have been focused on targeting enzyme active sites, through substrate, transition-state and even product analogues.[53,54,69,70] It has been hypothesised that drugs, which target allosteric sites lead to higher specificity than orthosteric drugs, which target an enzyme's functional site, as regions outside the active site tend to be less conserved among enzyme families. Allosteric drugs can act to modulate activity in a pathway, rather than eliminating activity altogether, and generally work when the enzyme is in its active state, which may be desirable in certain cases (i.e. when enzymatic activity is required by the cell).[85,95]

It remains unclear what features of the aromatic amino acids are required to achieve the allosteric response in this enzyme. Therefore, a detailed investigation into the binding of these aromatic amino acids and their ability to propagate allosteric signals to the active site, may aid future work into the targeting of this site for drug discovery.

An enzyme selects specific conformations of substrate that support the appropriate charge and the relative orientation of substrates to facilitate the desired reaction.[68,96] Active site reorganisation produced by binding of an allosteric effector can produce a response in catalytic rate. Type Iα DAH7PSs have been shown to undergo reorganisation of the active site in the presence of allosteric effector. In the Phe-sensitive isozyme from *E. coli,* the substrate, PEP was shown to adopt an alternative binding mode in the presence of Phe, which may be correlated to the reorganisation of crucial active site residues (PDB code: 1KFL).[22] It is hypothesised that *Nme*DAH7PS will undergo a similar structural reorganisation at the active site, as the active site reorganisation was also observed in the Tyr sensitive *Sce*DAH7PS (PDB code: 1OF6).

This chapter entails a detailed account of the binding and inhibition of *Nme*DAH7PS by the three aromatic amino acids, Phe, Tyr and Trp. The crystallisation of these ligands in the allosteric site of *Nme*DAH7PS depicts the reasons for selectivity between the three aromatic

amino acids, as well as similar structural reorganisation of the active site residues. These studies reveal that in the presence of allosteric ligand, the active site becomes pre-organised for the binding of the substrate PEP, which adopts an alternative binding mode and is less likely to react with E4P due to the reorganisation of active site residues. These insights into the allosteric mechanism pose an interesting avenue for future drug discovery efforts and our understanding of allostery in the enzyme, *Nme*DAH7PS.

3.2 RESULTS

3.2.1 The effect of Phe, Trp and Tyr on *Nme*DAH7PS activity.

The allosteric response of *Nme*DAH7PS to the three aromatic amino acids (Phe, Tyr and Trp) was tested using standard assay conditions (section 7.1.27). The results presented are similar to those of previously published work on this enzyme and illustrate that *Nme*DAH7PS is most sensitive to inhibition by Phe, is somewhat inhibited by Trp and least inhibited by Tyr (Figure 3.2.1.1).[26] The concentrations of metabolites in *E. coli* indicate the Tyr is the most abundant, followed by Phe and then Trp.[97] The heightened sensitivity of *Nme*DAH7PS to inhibition by Phe, suggests that regulation by Phe is sufficient to regulate the pathway in this organism, or that cellular levels of Tyr are regulated by other means.

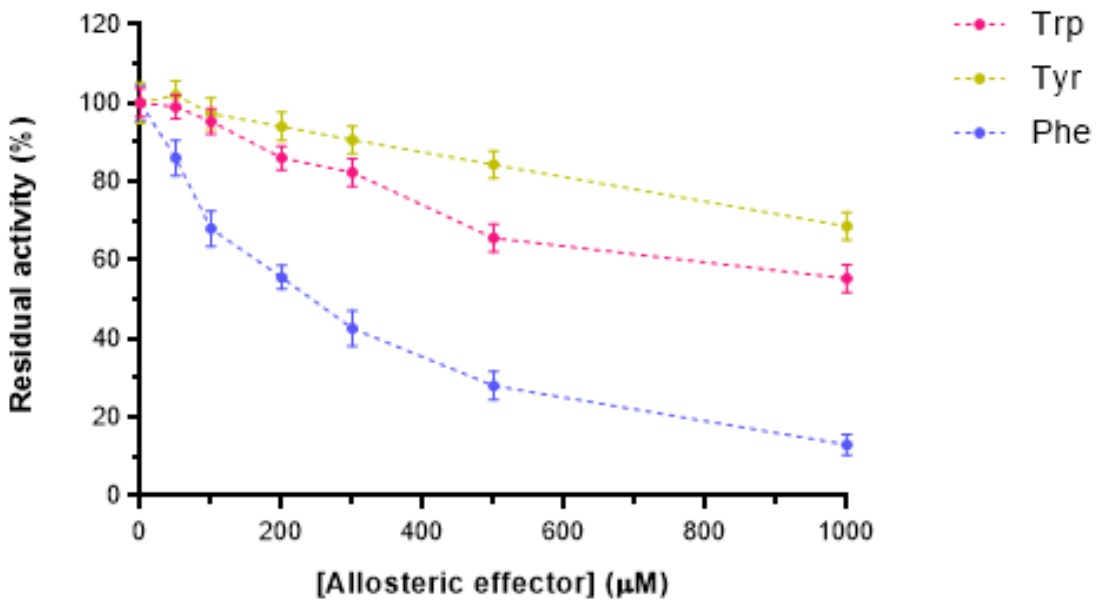


Figure 3.2.1.1 Allosteric effect of all three aromatic amino acids on the activity of NmeDAH7PS (2µL, 0.6 mg/mL) as a function of aromatic amino acid concentration, Tyrosine (green), Tryptophan (red) and Phenylalanine (blue). Concentrations of substrates were kept constant (PEP - 100 µM and E4P - 250 µM), while the concentration of the three aromatic amino acids was varied (50-1000 µM).

3.2.2 Binding of Phe, Trp and Tyr

Isothermal titration calorimetry (ITC) experiments were employed to determine the binding constants for the three aromatic amino acids. It was reported in the original published *Nme*DAH7PS data that binding of Tyr to *Nme*DAH7PS was not detected in the binding isotherm and therefore this titration was not attempted in this study.[26] Unfortunately, binding of Trp to *Nme*DAH7PS was too weak to fit an appropriate model to the binding isotherm after titration (Figure A.3.1). The binding of Phe to *Nme*DAH7PS was determined by titrating Phe into the enzyme and the binding isotherm was consistent with a two-site sequential binding model (Figure 4.2.2.1A).

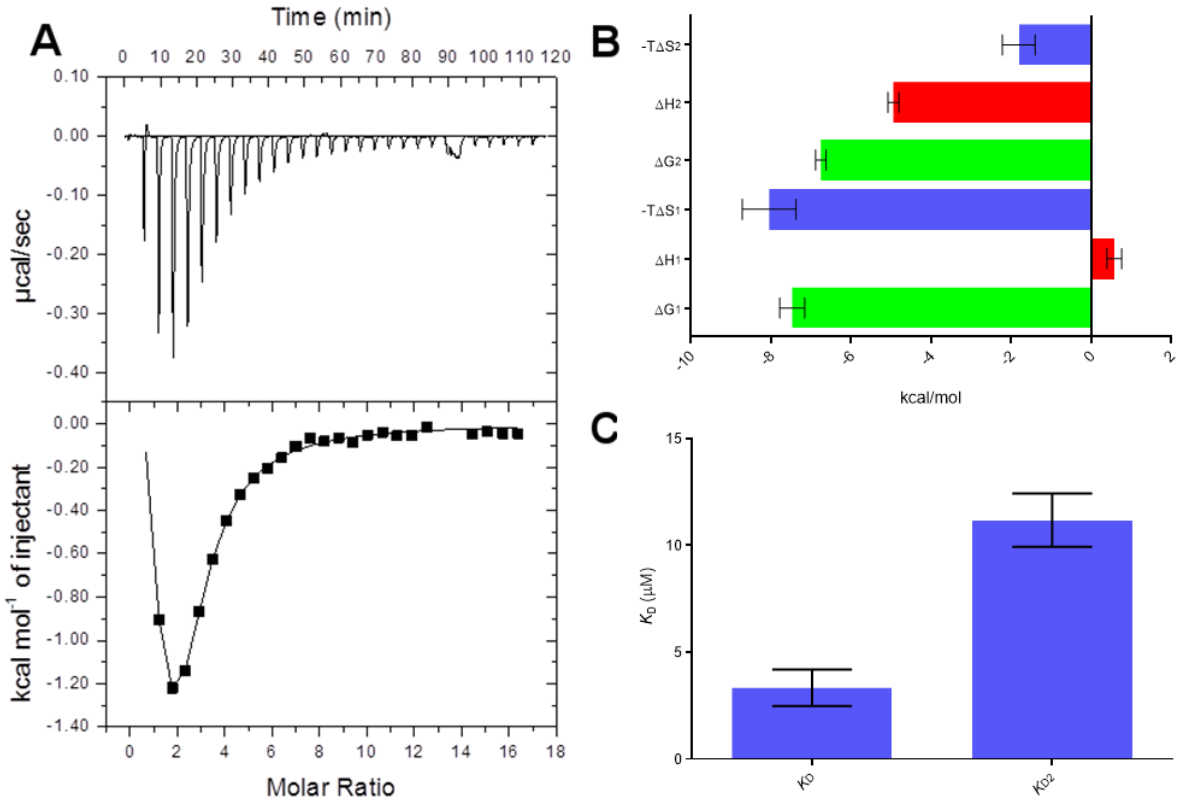


Figure 3.2.2.1 A) Raw data from ITC experiment involving the titration of 1.2 mM Phe into the cell containing 16 µM NmeDAH7PS corrected for heat of dilution. B) Comparison of the thermodynamic parameters (-TΔS-blue; ΔG-green and ΔH-red) for binding at both sites NmeDAH7PS. C) Comparison of dissociation constant (K_D) at both sites of the sequential binding model.

The binding of Phe to *Nme*DAH7PS had a K_D of 3 ± 1 µM at the first binding site and a K_D of 11 ± 1 µM for the second site. This indicates that binding occurs in an un-cooperative manner, where binding at the first available site decreases binding affinity at the second allosteric site.

The thermodynamic parameters obtained from the binding isotherm indicate that the large TΔS contribution to binding at the first site is due to the accessibility to two sites. The binding at this first site appears to be entropically driven whereas binding at the second site also has a large enthalpic contribution (Figure 3.2.2.1). Taken together, it is evident that binding of Phe to *Nme*DAH7PS is predominantly an entropically driven process.

3.2.3 Crystallographic characterisation of aromatic amino acid binding

In order to elucidate the binding mode of the aromatic amino acids, crystal structures of these ligands in complex with *Nme*DAH7PS were determined by X-ray diffraction methods. Crystals of *Nme*DAH7PS were grown in the absence of allosteric ligand using methods outlined in (section 7.1.30). Crystals were either transferred to droplets containing the crystallisation condition and the corresponding aromatic amino acid, or the crystallisation condition with aromatic amino acids was added directly to the drop containing the crystal in order to preserve crystal quality during soaking (section 7.1.33). Crystals were left to soak for 2-6 hours, where they were cryo-protected and flash frozen in liquid nitrogen prior to analysis (Data collection and refinement statistics are shown in Table 3.2.3.1).

Concentrations of amino acids used to obtain crystal structures containing these ligands varied considerably due to their vastly different affinity for the allosteric binding site of *Nme*DAH7PS. The final concentration of Phe, Trp and Tyr in the crystallisation drop used for soaking was 1 mM, 10 mM and 20 mM respectively.

It was apparent that none of the crystal structures contained electron density from the $|F_o\text{-}F_c|$ maps to indicate that the active site ligand, PEP was bound, even at low occupancy. This may have been lost during soaking experiments as PEP was in the purification buffer, but not in the crystallisation condition used for soaking. This remained an anomaly as previous structures of other type Iα DAH7PSs indicated a flipped binding mode for PEP.[28]

Instead of using soaking to diffuse the ligand into the crystal, attempts were made to co-crystallise the enzyme in the presence of (1 mM PEP and 1.5 mM Phe). Co-crystallisation of this protein in the presence of Phe is problematic in *Nme*DAH7PS due to the altered

conformation of the allosteric binding loop as it becomes ordered upon binding of the allosteric ligand. However, the use of several rounds of micro-seeding, generated crystals, that were of appropriate quality for analysis. Fortunately the structure was refined, and both ligands (Phe and PEP) were present at full occupancy (data collection and refinement statistics are given in Table 3.2.3.1).

NmeDAH7PS				
	L-Phenylalanine	**L-Phenylalanine, PEP**	**L-Tyrosine**	**L-Tryptophan**
Data Collection				
Crystal system; space group	Monoclinic, $P12_11$	Monoclinic, $P12_11$	Monoclinic, $P12_11$	Monoclinic, $P12_11$
Unit cell parameters				
a, b, c (Å)	73.61, 143.50, 75.18	75.46, 140.31, 75.50	72.94, 141.30, 74.68	73.43, 138.04, 75.87
α, β, γ (º)	90, 96.18, 90	90, 96.89, 90	90, 96.28, 90	90, 96.58, 90
Resolution range (Å)	47.82-2.19 (2.23-2.19)	74.96-2.05 (2.09-2.05)	74.23-2.31 (2.37-2.31)	75.37-2.23 (2.28-2.23)
Measurements	301997	364904	492196	546860
Unique reflections	79522	97425	65792	72956
Redundancy	3.8 (3.8)	3.7 (3.7)	7.5 (7.7)	7.5 (7.5)
Completeness (%)	99.9 (100.0)	99.9 (99.9)	100.0 (99.9)	99.9 (100.0)
I/σ (I)	9.3 (1.2)	8.3 (1.82)	18.1 (3.75)	14.6 (2.26)
R_{merge}	0.084 (1.05)	0.109 (0.686)	0.072 (0.499)	0.086 (1.10)
$CC_{1/2}$	0.50	0.50	0.92	0.66
Wilson *B* value $(Å)^2$	39.81	23.9	37.4	43.2
Matthews coefficient	2.60 (52.65)	2.61 (52.91)	2.52 (51.15)	2.51 (51.09)
Refinement				
R_{work}	0.1890	0.1687	0.1938	0.1915
R_{free}	0.2357	0.2120	0.2387	0.2378
Chain length	351	351	351	351
Observed number of residues	345 (Chain A&B), 345 (Chain C&D)	345 (Chain A &B), 347 (Chain C), 344 (chain D)	346 (In all chains)	345 (Chain C), 348 (Chain D), 344 (Chain A), 342 (Chain B)
Water molecules	330	778	290	216
Other	9 (Mn^{2+}, SO_4^{2-}, PEG)	11 (Mn^{2+}, SO_4^{2-}, PEG)	9 (Mn^{2+}, SO_4^{2-}, TMO)	4 (Mn^{2+})
Ligand	4	8	4	4
Mean *B* $(Å)^2$				
Protein	36.47	28.34	38.37	34.10
Water	43.82	30.05	38.05	49.51
Other	67.46	48.77	48.41	60.35
Ligand	37.13	30.97 (PEP), 20.92 (Phe)	38.11	57.92
R.m.s.d from target values				
Bond lengths (Å)	0.0095	0.0094	0.0102	0.0092
Bond angles	1.2712	1.3263	1.3477	1.2624
Dihedral angles	0.0690	0.0768	0.0758	0.0718
Ramachandran				
Preferred (%)	97.69	97.25	97.45	97.51
Allowed (%)	1.73	2.61	1.97	1.90
Outliers (%)	0.58	0.14	0.58	0.59
PDB Entry	4UC5	5DCB	5DCD	5DCE

Table 3.2.3.1 Data collection and refinement statistics for the four ligand-bound structures of NmeDAH7PS.

3.2.3.1 Binding modes of Phe, Trp and Tyr at NmeDAH7PS allosteric site

All liganded structures crystallised in a monoclinic space group ($P12_11$), as that of the PEP-bound *Nme*DAH7PS^{WT} (PDB code: 4HSN). The Phe, Phe + PEP, Tyr and Trp bound structures were refined to resolutions of 2.19, 2.05, 2.31 and 2.23 Å respectively, and all ligands have been modelled into their respective structures at full occupancy.

The refined structures all exhibit similar structural characteristics. The RMSD for the alignment of all Cα atoms of the tight dimers of the Phe-bound structure (PDB code: 4UC5) with the PEP-bound structure (PDB code: 4HSN) was 0.583-0.752 Å. This is indicative of very little conformational change upon binding of the most potent allosteric effector. The most severely affected regions of the enzyme in the presence of Phe are localised to the $\beta_2\alpha_2$ loop, the $\beta_3\alpha_3$ loop, α_3 helix and the β-hairpin insertion (Figure 3.2.3.1.1). The variation seen in these regions is consistent throughout all liganded structures.

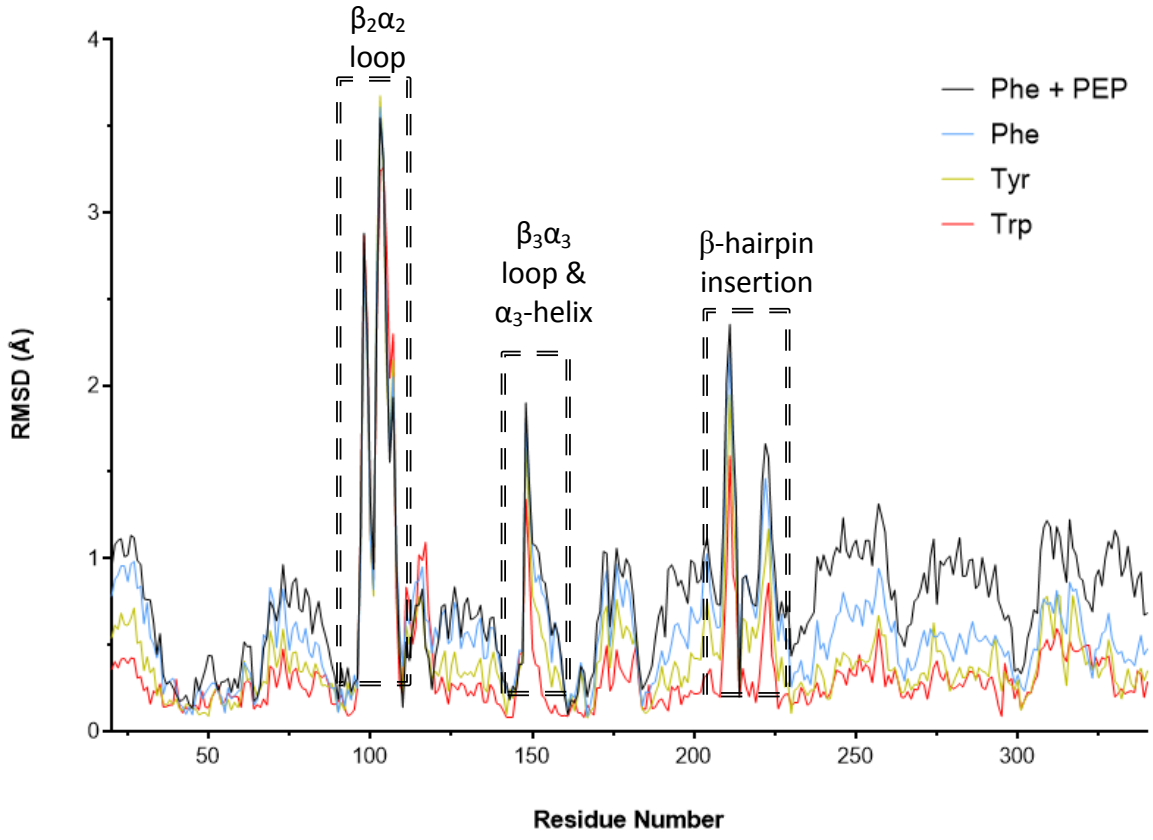


Figure 3.2.3.1.1. Root mean squared deviation plot of all Cα-atoms per residue of ligand-bound structures of NmeDAH7PS illustrating important regions of variation are highlighted by a black double dashed box. Liganded structures against PEP-bound structure (PDB code: 4HSN). Trp (red), Tyr (gold), Phe (blue) and Phe + PEP (black)

The Trp- and Tyr-bound structures (PDB code: 5DCD and 5DCE) have very similar structural characteristics to the Phe-bound structure. They both show significant variation from the PEP-bound structure in the same regions ($\beta_2\alpha_2$ loop, the $\beta_3\alpha_3$ loop and the β-hairpin insertion) (Figure 3.2.3.1.1). Alignment of the tight dimers of the Tyr-bound structure (PDB code: 5DCD) with that of the Phe-bound structure (PDB code 4UC5) yielded an RMSD of 0.220-0.399 Å. This result is indicative of very small variation in the presence of Tyr and implies a very similar structural reorganisation upon ligand binding to *Nme*DAH7PS. The Trp bound structure (PDB code: 5DCE) shows a slightly greater degree of variation from the Phe-bound structure and has an RMSD for the alignment of the tight dimers of 0.460-0.572 Å. This structural variation is localised to the aforementioned regions that are most affected in the presence of allosteric effector (Figure 3.2.3.1.1 and 3.2.3.1.2).

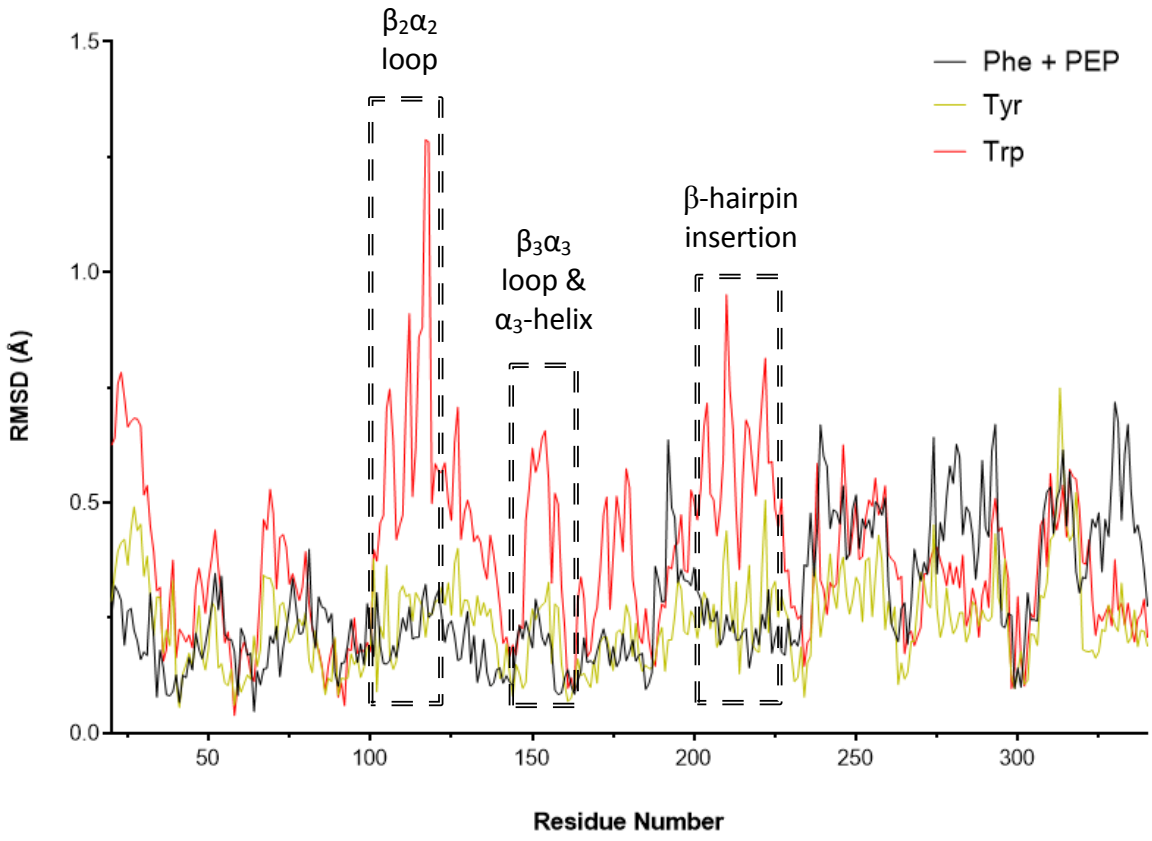


Figure 3.2.3.1.2. Root mean squared deviation plot all Cα-atoms per residue of ligand-bound structures of NmeDAH7PS illustrating important regions of variation are highlighted by a black double dashed box. Liganded structures against Phe-bound structure (PDB code: 4UC5). Trp (red), Tyr (gold), and Phe + PEP (black).

Unsurprisingly, the Phe+PEP-bound structure (PDB code: 5DCB) was very similar to the Phe-bound structure (PDB code: 4UC5) with an RMSD for the alignment of the tight dimers of

0.237-0.326 Å. However, there are small differences in the positioning of the $\beta_7\alpha_7$, $\beta_8\alpha_8$ loops and the α_8 helix located towards the C-terminus of the enzyme (res. 260-350).

All three aromatic amino acids adopt a very similar binding mode at the allosteric site of *Nme*DAH7PS. Tyr and Phe are present at full occupancy in the allosteric site of *Nme*DAH7PS, whereas, Trp is present at less than full occupancy. The lower occupancy of Trp is evidenced by slightly higher temperature factors for this ligand and N-terminal tail residues compared to central residues. The conformations of the side-chains of surrounding residues are nearly identical when all ligands are superimposed. The amino-carboxylate moiety of Phe, Trp and Tyr are almost identically positioned within the binding pocket, however, the side-chains of these ligands interact in a slightly different manner with residues of the binding site. It is likely that favourable and unfavourable interactions within the pocket are the driving force behind selectivity in *Nme*DAH7PS (Figure 3.2.3.1.3).

The selectivity for aromatic amino acids in *Nme*DAH7PS likely arises from the largely hydrophobic pocket as well as a highly specific and conserved N-terminal extension (allosteric binding loop). The residues deep within the pocket interact with the aromatic rings of the bound ligands, whereas the residues of the loop and other polar residues interact with the amine and carboxylate moiety of the aromatic amino acid (Table 3.2.3.1.1 and Figure 3.2.3.1.3).

Liganded structure	***Lig-COO^- Asp9 backbone (NH2)***	***Lig-COO^- Ser182 backbone (NH2)***	***Lig-NH_3^+ Asp8 side-chain (C=O)***	***Lig-NH_3^+ Asp9 side-chain (C=O)***	***Lig- NH_3^+ Gln153 side-chain (C=O)***	***Lig- π Phe211 side-chain (π)***	***Lig-Ser213 side-chain (OH)***	***Lig-NH_3^+ to HOH***
Phe	2.9 Å	2.8 Å	3.5 Å	2.8 Å	2.9 Å	3.9 Å	3.2 Å	2.9 Å
Phe+PEP	2.8 Å	2.8 Å	3.5 Å	2.8 Å	2.9 Å	3.8 Å	3.4 Å	2.7 Å
Tyr	2.9 Å	2.7 Å	3.7 Å	2.8 Å	2.9 Å	4.1 Å	2.4 Å	2.7 Å
Trp	2.9 Å	2.8 Å	3.7 Å	2.9 Å	2.8 Å	4.4 Å	3.5 Å	N/A

Table 3.2.3.1.1 Key protein-ligand interactions in the Phe, Tyr, Trp and Phe+PEP-bound structures. Green-more favourable interactions between structures. Red- unfavourable interactions.

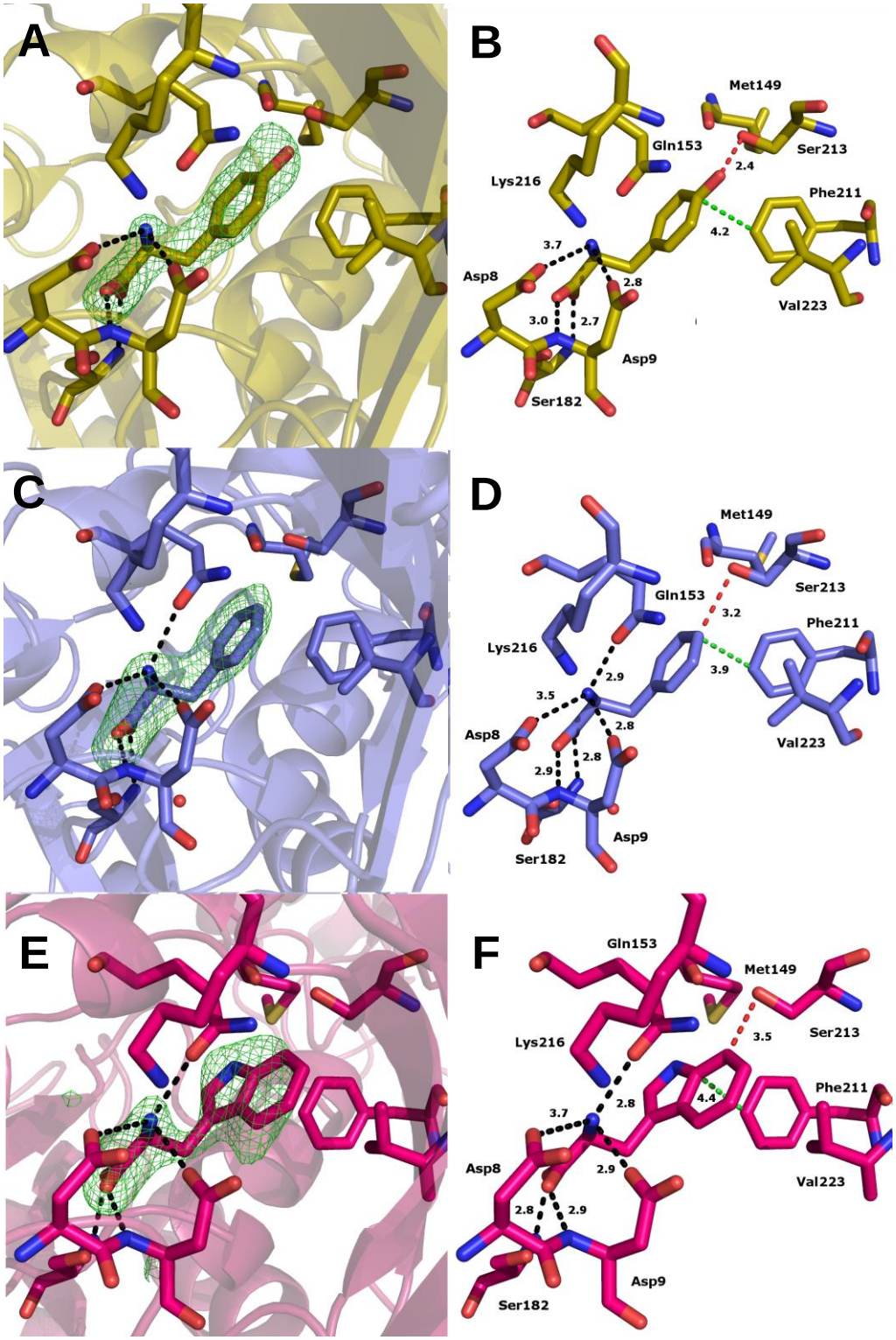


Figure 3.2.3.1.3 Binding modes of the three aromatic amino acids (Phe, Trp and Tyr). Hydrogen bonds are indicated in black. Distance of π-interaction with Phe211 indicated in green and an unfavourable interaction with Ser213 indicated in red. A) $|F_o-F_c|$ map of the Tyr binding site from chain A of Tyr-bound structure (PDB code: 5DCD) contoured at 3σ. B) The Tyr binding site with average interaction distances from all chains depicted. C) $|F_o-F_c|$ map of the Phe binding site from chain A of Phe-bound structure (PDB code: 4UC5) contoured at 3σ. D) The Phe binding site with average interaction distances from all chains depicted. E) $|F_o-F_c|$ map of the Trp binding site from chain A of Phe-bound structure (PDB code: 5DCE) contoured at 3σ. F) The Trp binding site with average interaction distances from all chains depicted.

Residues (4-10) of the N-terminal allosteric binding loop become ordered upon binding of the aromatic amino acid. The side-chain of Asp9 forms a hydrogen bond with the amine of the bound ligands and the backbone peptide amino group of Asp9 forms a hydrogen bond with the carboxylate of the bound ligands. Interestingly, the distance between the amine of the bound ligand and the side-chain carboxylate of Asp8 is closer (~0.2 Å) in the Phe-bound structure than the Trp and Tyr bound structures and may reflect a tighter interaction of Phe with this loop.

Inspection of the interactions of the ligand with residues Phe211 and Ser213, located deep within the site, reveals some interesting and noteworthy differences. The recruitment of Phe211 and repulsion of Ser213 are believed to be very important for ligand binding and allosteric communication.[22,26] Comparison of the distance for the edge to face π-interaction between Phe211 and the bound ligand reveal some subtle differences between the binding modes of the aromatic amino acids at the allosteric site. In the Phe-bound structures, this distance is as close as 3.8 Å, whereas, in the Tyr bound structure the distance is 4.2 Å and in the Trp bound structure the distance gets as large as 4.5 Å.

Ser213 does not form a hydrogen bond with any of the bound ligands. In the Phe and Trp bound structures, the distance between the side-chain hydroxyl group of Ser213 and the bound ligand is >3.2 Å. However, in the Tyr bound structure, the hydroxyl group of the Tyr bound ligand is ~2.4 Å away from the side-chain hydroxyl of Ser213 (Table 3.2.3.1.1). At this close proximity, electrostatic and Van der Waals interactions become repulsive. This is consistent with a significant decrease in affinity of the binding site for Tyr and may explain the increase in distance from Phe211. Ser213 has been shown to play a crucial role in selectivity within the binding site of *Nme*DAH7PS and this is probably due to the unfavourable interaction with the hydroxyl functionality of Tyr.[26] The location of residues, Phe211 and Ser213 on the β_{6a} sheet at the tight-dimer interface and the interaction between this interface and the $\beta_2\alpha_2$ loop means that interactions here may have implications for catalysis.

3.2.3.2 Phe, Trp and Tyr cause similar active site reorganisation

The $\beta_2\alpha_2$ loop contains the KPR(T/S) motif, which is important for catalysis and is intimately associated with residues of the adjacent monomer across the tight dimer interface.[40,42,54,98] The RMSD plots per residue (Figure 3.2.3.1) indicate that there is a new conformation adopted by the $\beta_2\alpha_2$ loop in the presence of the allosteric ligands (Phe, Trp or Tyr) and that this new conformation is consistent for all structures containing an allosteric ligand (Figure 3.2.3.2.1).

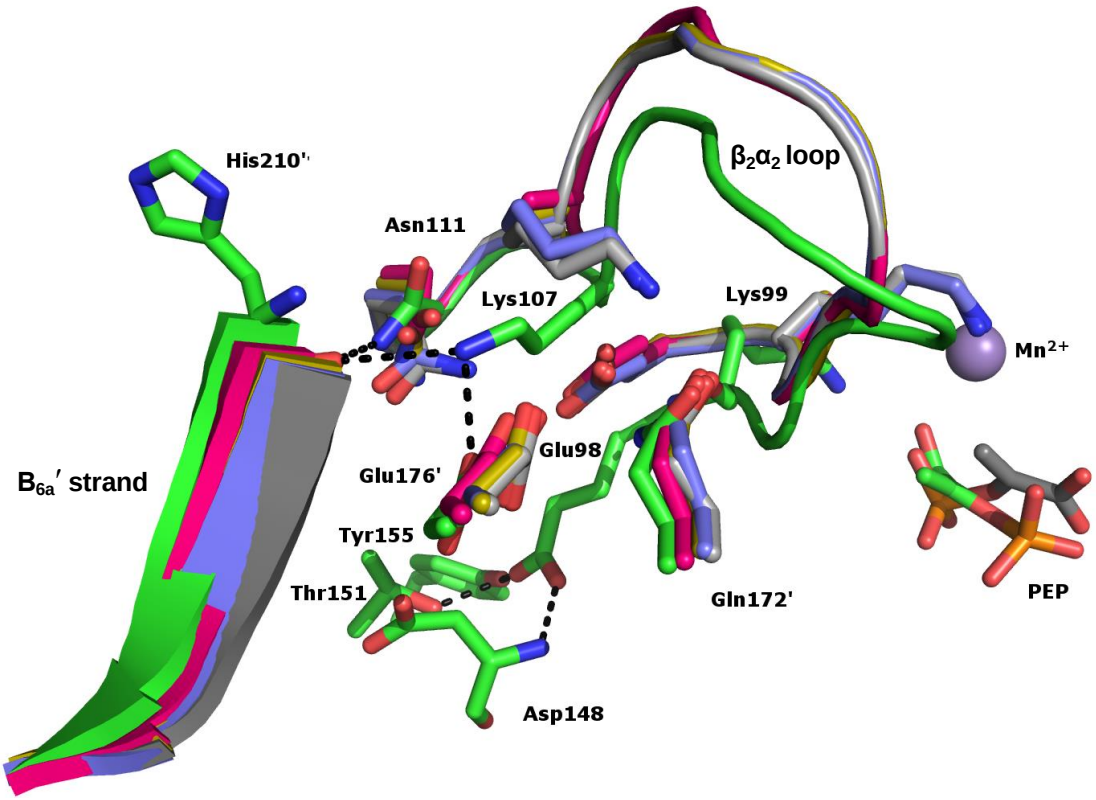


Figure 3.2.3.2.1 Overlay of the Phe-, Trp-, Tyr-, Phe + PEP- and PEP-bound NmeDAH7PS structures indicating the new conformation of $\beta_2\alpha_2$ loop and the loss of hydrogen bonds upon binding of allosteric effector. PEP-bound structure (green), Phe-bound structure (blue), Trp bound structure (red), Tyr bound structure (gold) and Phe+PEP-bound structure (grey). Hydrogen bonds broken upon ligand binding are indicated in black.

At the variable concentrations at which these ligands were soaked into the enzyme, the structural effect on the $\beta_2\alpha_2$ loop is analogous. This implies that binding affinity is the main contributor to the decreased inhibition observed for Tyr and Trp. The new conformation of the $\beta_2\alpha_2$ loop is a result of three main structural changes. There is the loss of a hydrogen bond between Glu98 and residues, Asp148, Thr151 and Tyr155 of the $\beta_3\alpha_3$ loop and α_3 helix. Lys107 loses hydrogen bonds across the tight dimer interface with His210′ and Glu176′. Finally, the backbone carbonyl of the absolutely conserved Lys99 no longer forms a hydrogen bond with

the side-chain of Gln172 from the adjacent chain. These three changes, which are observed for all liganded structures contribute to the new conformation of $\beta_2\alpha_2$ loop and may have significant implications for catalysis.

3.2.3.3 Flipped binding mode of PEP caused by pre-organised active site

Previous investigations into the Phe sensitive type Iα DAH7PS enzyme from *E. coli* indicated that PEP adopts a flipped binding mode at the active site in the presence of allosteric ligand (PDB code: 1KFL).[22] What causes PEP to adopt this flipped binding mode and whether or not it adopts this binding mode in the regulated form of *Nme*DAH7PS remained unclear. A crystal structure was subsequently obtained via co-crystallisation of both PEP and Phe with *Nme*DAH7PS. Co-crystallisation of PEP with Tyr and Trp was not attempted due to time constraints and the very high concentrations required for ligand binding. The PEP+Phe-bound structure (PDB code: 5DCB) was refined to 2.05 Å and showed evidence for both ligands bound at full occupancy at their respective sites. Interestingly, like that of the *Eco*DAH7PS structure, PEP adopted a flipped orientation in the active site (Figure 3.2.3.3.1).

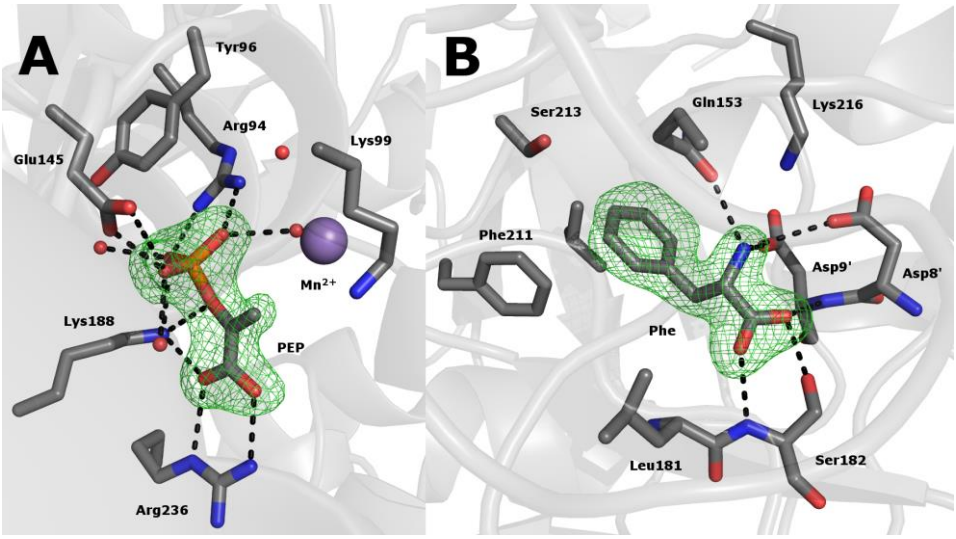


Figure 3.2.3.3.1 A) Active site of the Phe+PEP-bound NmeDAH7PS (PDB code: 5DCB), indicating the presence of PEP at full occupancy. $|F_o\text{-}F_c|$ map contoured at 3 σ. B) Allosteric site of the Phe+PEP-bound NmeDAH7PS (PDB code: 5DCB), indicating the presence of Phe at full occupancy. $|F_o\text{-}F_c|$ map contoured at 3 σ.

Comparison of the binding mode in the absence (PDB code: 4HSN) and presence (PDB code: 5DCB) of Phe indicated some interesting differences. The dihedral angle between the carboxylate and the C_2=C_3 of PEP is significantly reduced in the presence of Phe from 44 ± 2° to 15 ± 1° (Figure 3.2.3.3.2). This is hypothesised from *ab initio* quantum chemistry

calculations to reduce the reactivity of the enol moiety of the substrate by reducing the charge build-up at C_3 of PEP.[99] Interestingly the network of water molecules remains similar despite the new binding mode of PEP. However, the nucleophilic water (WAT1) identified from previous studies based on its proximity to C_2 of PEP or inhibitor has moved.[100] The nearest water molecule to C_2 of PEP (WAT1) is now located on the *si* face of PEP in the presence of Phe.

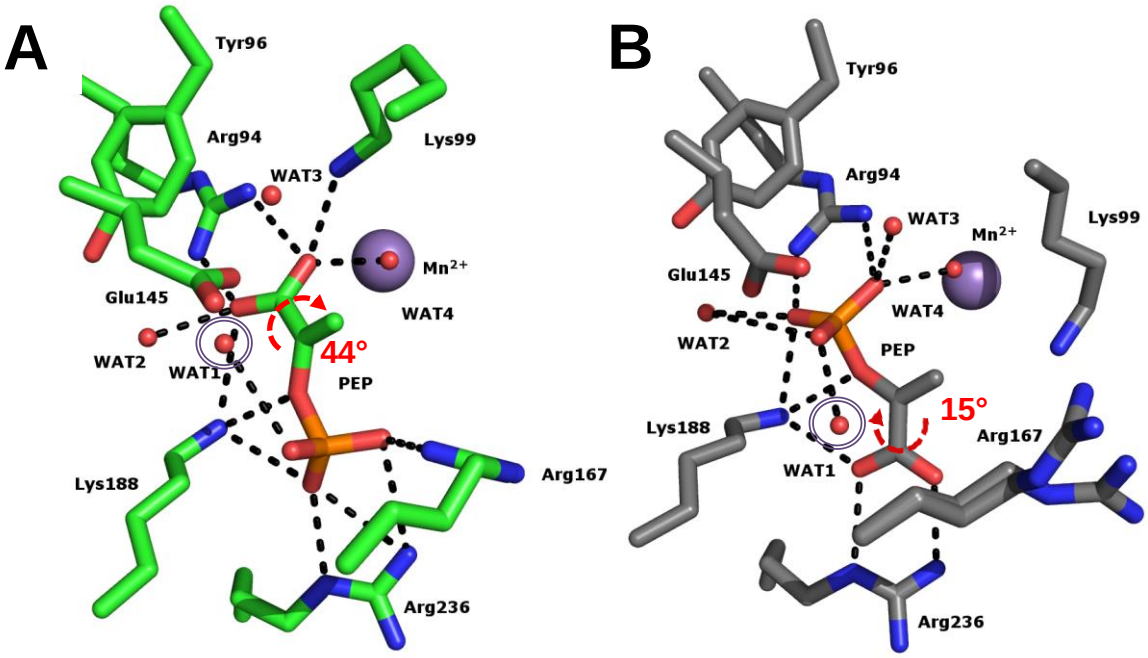


Figure 3.2.3.3.2 A) All bonding interactions between PEP and active site from active enzyme (PDB code: 4HSN) and the dihedral angle between the carboxylate and the C=C of PEP B) All bonding interactions between PEP and active site from Phe-bound enzyme (PDB code: 5DCB) and the dihedral angle between the carboxylate and the C=C of PEP. Proposed nucleophilic water(WAT1) indicated by a double lined circle.

The movement of Lys99 in the presence of Phe is conserved for all structures containing an allosteric ligand. This new position, adopted by Lys99 is correlated to the new conformation of the $\beta_2\alpha_2$ loop and enables the formation of a salt bridge with the metal binding residue, Asp324. Another noticeable difference at the active site is the new conformation of Arg167 from the α_4 helix that can no longer form the salt bridge with the phosphate of PEP observed in the PEP-bound structure (PDB code: 4HSN) (Figure 3.2.3.3.2). The new conformation of these two residues makes the approach of PEP at the active site significantly different and may contribute to the new binding mode.

Comparison of the active site residues responsible for both PEP and E4P binding from the Phe-bound (PDB code: 4UC5) and Phe+PEP-bound (PDB code: 5DCE) indicate no difference in their positioning (Figure 3.2.3.3.3). The similarity between the conformations and positioning of side-chains of the active site residues in these two structures regardless of the presence of PEP indicates that the active site is pre-organised for PEP binding in its flipped orientation in the presence of Phe.

The new conformation of the active site residues also has implications for the other substrate E4P. During catalysis, the aldehyde of E4P requires activation as an electrophile for nucleophilic attack by the enol moiety of PEP.[32,49,51,53,63] In the presence of allosteric ligand, the formation of the salt bridge between Lys99, part of the KPR(T/S) motif, and Asp324 structurally hinders access of the aldehyde of E4P to the divalent metal ion. Furthermore, Arg101 from the conserved KPR(T/S) motif located on the $\beta_2\alpha_2$ loop, responsible for binding the phosphate of E4P[33,63] adopted a different conformation in all allosteric ligand bound *Nme*DAH7PS structures compared to *Eco* and *Sce*DAH7PS (Figure 3.2.3.3.3). In the Phe-bound form of *Eco*DAH7PS, the equivalent Arg residue forms a salt bridge with Glu96. This has also been observed in two out of eight chains of the Tyr-regulated *Sce*DAH7PS structure (PDB code: 1OF6). However, in *Nme*DAH7PS, Arg101 does not interact with Glu98, regardless of the presence or absence of PEP at the active site. Instead it adopts a similar conformation to what is observed in the PEP-bound structure (PDB code: 4HSN). The reorganisation of the $\beta_2\alpha_2$ loop and surrounding residues in *Nme*DAH7PS may be responsible for this unique conformation of Arg101. In particular, the position of Lys107 and its interaction with Gln172 and Glu98 in the Phe-bound *Nme*DAH7PS structures may limit the ability of Arg101 to form a salt bridge with Glu98. The variability in conformation Arg from the KPR(T/S) motif may contribute to the increased sensitivity of *Eco*DAH7PS to concentrations of Phe reported previously.[101]

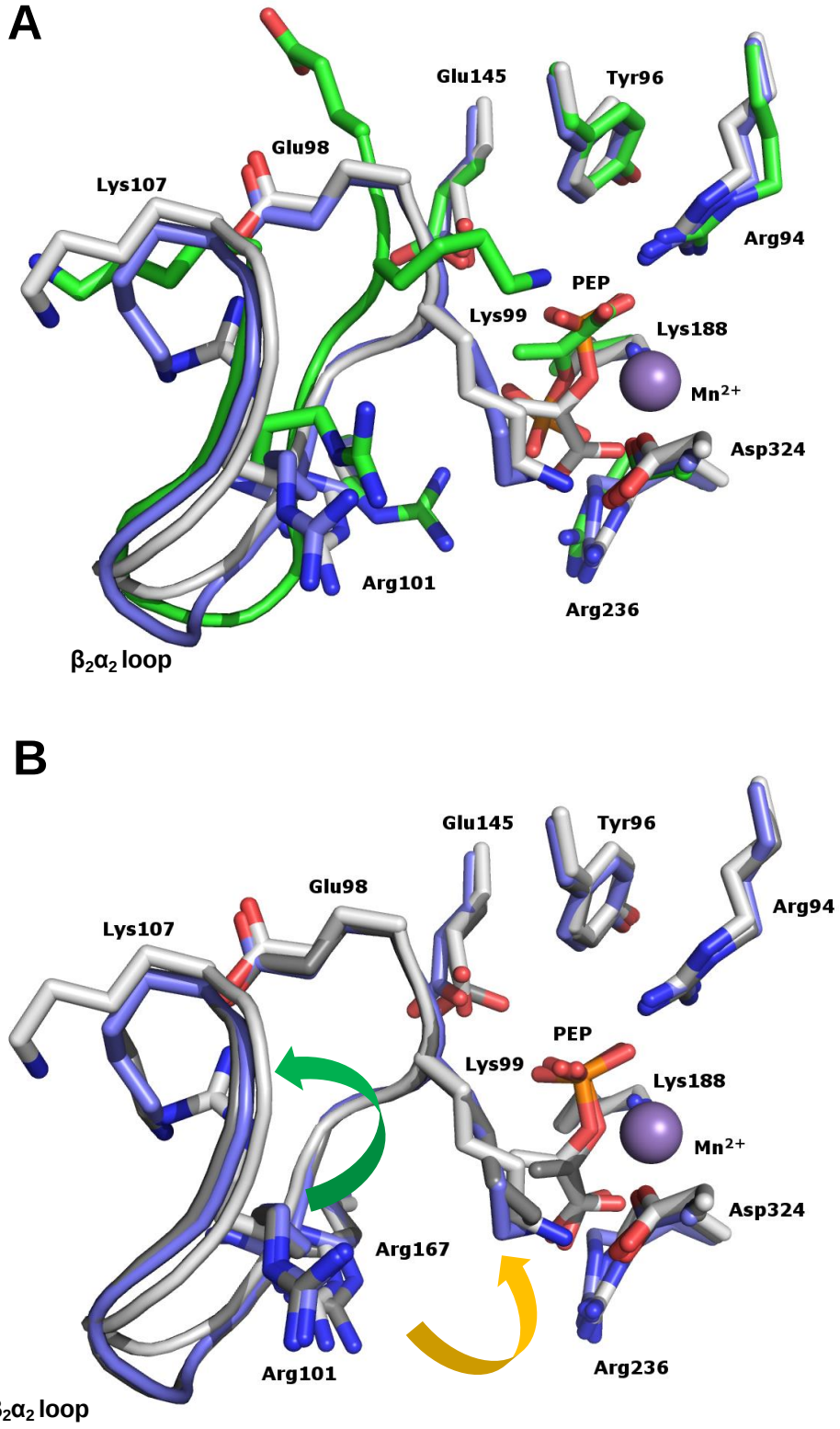


Figure 3.2.3.3.3 Active site of type Iα DAH7PS. PEP-bound NmeDAH7PS, PDB code: 4HSN (green), PEP and Phe-bound EcoDAH7PS, PDB code: 1KFL (white), Phe-bound NmeDAH7PS, PDB code: 4UC5 (blue), PEP+Phe-bound NmeDAH7PS, PDB code: 5DCB (dark grey). A) Overlay of PEP binding site from EcoDAH7PS (PEP+Phe-bound), NmeDAH7PS (PEP-bound) and NmeDAH7PS (Phe-bound) illustrating the flipped binding mode of PEP, movement of catalytic Lys99 residue and overall structural conservation at the site. B) Overlay of EcoDAH7PS (PEP+Phe-bound), NmeDAH7PS (Phe-bound) and NmeDAH7PS (Phe+PEP-bound). Yellow arrow indicates channel for PEP binding as Lys99 partially gates the site. Green arrow indicates movement of Arg residue.

3.3 DISCUSSION

An understanding of the molecular details that determine the selectivity of *Nme*DAH7PS to inhibition by the three aromatic amino acids was not well understood. It was found that the enzyme is most sensitive to Phe, which was in accordance with previous results, and indicates that the other aromatic amino acids have little effect on the enzyme at the likely cellular concentrations of these metabolites.[97] The ITC experiments revealed a negative cooperative mechanism of binding Phe from the fit of a two-site sequential model to the binding isotherm. However, Trp and Tyr binding to the enzyme was not detectable by this experimental method, even though some inhibition of catalytic activity is observed at high concentrations.

To aid in the elucidation of mechanism of allosteric inhibition, X-ray crystallographic data with these ligands bound at the allosteric site were sought. In these structures Tyr and Trp adopt a very similar binding mode to that observed for Phe at the allosteric site of *Nme*DAH7PS. The interactions of residues Asp8, Asp9, Gln153 and Ser182 in with the amino-carboxylate portion of the aromatic amino acid are nearly identical for all ligands. However, there are subtle differences in the interactions with residues, Phe211 and Ser213 deep within the pocket, which may explain the variation in sensitivity observed. The binding of Tyr in the site is unfavourable due to a repulsive interaction with the hydroxyl moiety of Ser213. This is obviously an important feature, as mutation of this residue to glycine has been shown to specifically change selectivity of this site to Tyr, therefore, it is not surprising that it is pivotal to ligand selectivity.[26] Phe forms the strongest π-interaction with Phe211 at a distance of ~3.8 Å and Trp forms the weakest interaction (~4.7 Å) and this may correlate to this increased affinity of the site for Phe. However, It is difficult to correlate differences in binding mode to binding affinity as Trp and Tyr binding cannot be observed by ITC. Therefore, it is uncertain whether these structural differences affect binding affinity or allosteric response, as inhibition at the same concentrations of aromatic amino acids varies considerably, but this cannot be correlated to the binding mode of the respective aromatic amino acid.

All three aromatic amino acids have the same effect on the structural reorganisation of the PEP and E4P ($\beta_2\alpha_2$ loop) binding site, which was shown to adopt the same conformation in the presence and absence of PEP. This similarity implies that the PEP binding site is pre-organised

for the binding of PEP in its flipped binding mode, regardless of what aromatic amino acid is bound at the allosteric site. The formation of a salt bridge between Lys99 and Asp324 may contribute to flipped binding mode of PEP.

The new conformation of this site reduces the reactivity of both substrates. PEP reactivity is reduced through a reduction in the dihedral angle between carboxylate and enol functionality, thus reducing the accumulation of charge due to sp3 hybridisation at the C_3 position of PEP, ultimately decreasing its ability to act as a nucleophile during catalysis.[99]. Furthermore, the salt bridge blocks the access of E4P to the divalent metal ion, which is crucial for activation and catalysis. Remarkably, Arg101, which is responsible for binding the phosphate of E4P remains in the same conformation as is seen in the PEP-bound structure. This is contrary to what is observed for other type Iα DAH7PSs, where this residue forms a salt bridge with the equivalent residue to Glu98 in the presence of allosteric inhibitor.

As sensitivity to regulation by Phe in *Nme*DAH7PS is reduced, compared to *Eco*DAH7PS (~30% at 100 μM Phe), perhaps it is not surprising that we see negative cooperativity in the binding isotherm of *Nme*DAH7PS from ITC. Whereby, occupation of one binding site decreases the binding affinity at the remaining binding site. Perhaps the increased sensitivity seen in *Eco*DAH7PS is a direct result of the ability to form this salt-bridge. The exact reasons for the formation of this salt bridge are difficult to ascertain, but imply that dynamics are very important for the conformations adopted by the $\beta_2\alpha_2$ loop.

The combination of regulatory features of *Nme*DAH7PS outlined above, make it unique, not only compared to other enzymes, but even to other characterised type Iα DAH7PS. This has interesting implications for the evolution of allostery in DAH7PS that are specific to the cellular environment and genome of the individual organism.

CHAPTER 4- IMPORTANCE OF INTERACTIONS AT THE TETRAMER INTERFACE OF *NME*DAH7PS

CHAPTER 4 PART I-The functional unit of NmeDAH7PS is dimeric

4.1 INTRODUCTION

Previous work in this area involved the mutation of a residue (E24Q) involved in a salt bridge interaction across the tetramer interface of *Eco*DAH7PS, yielding a dimeric form of the enzyme.[31] Unfortunately, there were neither solution state nor functional studies reported for this variant from *Eco*DAH7PS. This residue and salt-bridge interaction is conserved at the tetramer interface of *Nme*DAH7PS and is formed by the interaction between Glu27 and Arg126.

This study involved the construction of an Arg126Ser mutation at the tetramer interface to form a dimeric variant of *Nme*DAH7PS. This was done in order to identify the necessity for tetramerisation in *Nme*DAH7PS. The characterisation of this variant enzyme and comparison to *Nme*DAH7PS^{WT} revealed remarkable similarities upon dimerisation.

This drafted manuscript has been accepted for publication by PLoS ONE

Heyes, L. C.*, Cross, P. J.*, Zhang, S.; Nazmi, A. R. & Parker, E. J. The functional unit of *Neisseria meningitidis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase is dimeric. *PLoS ONE* **XX,** XXX-XXX (XXXX).

Dr P. J. Cross and S. Zhang cloned, expressed and purified *Nme*DAH7PS^{R126S}. Dr P. J. Cross completed the kinetic characterisation and small angle x-ray scattering experiments and analysis. Dr A. R. Nazmi performed and analysed the analytical ultracentrifugation experiments. Prof E.J. Parker designed experiments and contributed to the writing of the manuscript. The author of this thesis wrote the publication manuscript and all co-authors contributed to the writing and/or revision of the manuscript. The author of this thesis carried out the time dependent loss of kinetic activity experiments, refinement and deposition of the crystal structure, the interface analysis, the analytical size-exclusion chromatography, the isothermal titration calorimetry and the differential scanning calorimetry experiments.

The functional unit of *Neisseria meningitidis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase is dimeric

Penelope J Cross[1@], *Logan C Heyes*[1@], *Shiwen Zhang*[1], *Ali Reza Nazmi*[1] *and Emily J Parker*[1*]

[1]Biomolecular Interaction Centre and Department of Chemistry, University of Canterbury, Christchurch, New Zealand

Short title: *N. meningitidis* DAH7PS quaternary structure

*Corresponding author. Department of Chemistry, University of Canterbury, Private Bag 4800, Christchurch, Telephone (+64) 3 364 5682, Fax (+64) 3 364 2110, Email: emily.parker@canterbury.ac.nz

[@]These authors contributed equally to this work

[‡]This research was funded, in part, by the New Zealand Marsden Fund (UOC1105).

ABBREVIATIONS

BTP, 1,3-bis[tris(hydroxymethyl)methylamino]propane; DAHPS, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate; DAH7PS, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase; E4P, erythrose 4-phosphate; PEP, phosphoenol pyruvate; Phe, phenylalanine; RMSD, root mean standard deviation; Trp, tryptophan; Tyr, tyrosine

ABSTRACT

Neisseria meningitidis 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (*Nme*DAH7PS) adopts a homotetrameric structure consisting of an extensive and a less extensive interface. Perturbation of the less extensive interface through a single mutation of a salt bridge (Arg126-Glu27) formed at the tetramer interface of all chains resulted in a dimeric DAH7PS in solution, as determined by small angle X-ray scattering, analytical ultracentrifugation and analytical size-exclusion chromatography. The dimeric *Nme*DAH7PS^{R126S} variant was shown to be catalytically active in the aldol-like condensation reaction between D-erythrose 4-phosphate and phosphoenolpyruvate (PEP), and allosterically inhibited by L-phenylalanine (Phe) to the same extent as the wild-type enzyme. The dimeric *Nme*DAH7PS^{R126S} variant exhibited a slight reduction in thermal stability by differential scanning calorimetry experiments and a slow loss of activity over time compared to the wild-type enzyme. Although *Nme*DAH7PS^{R126S} crystallised as a tetramer, like the wild-type enzyme, structural asymmetry at the less extensive interface was observed consistent with its destabilisation. The tetrameric association enabled by this Arg126-Glu27 salt-bridge appears to contribute solely to the stability of the protein, ultimately revealing that the functional unit of *Nme*DAH7PS is dimeric.

Keywords:

Allostery, Shikimate pathway, DAHPS, DAH7PS, Quaternary structure

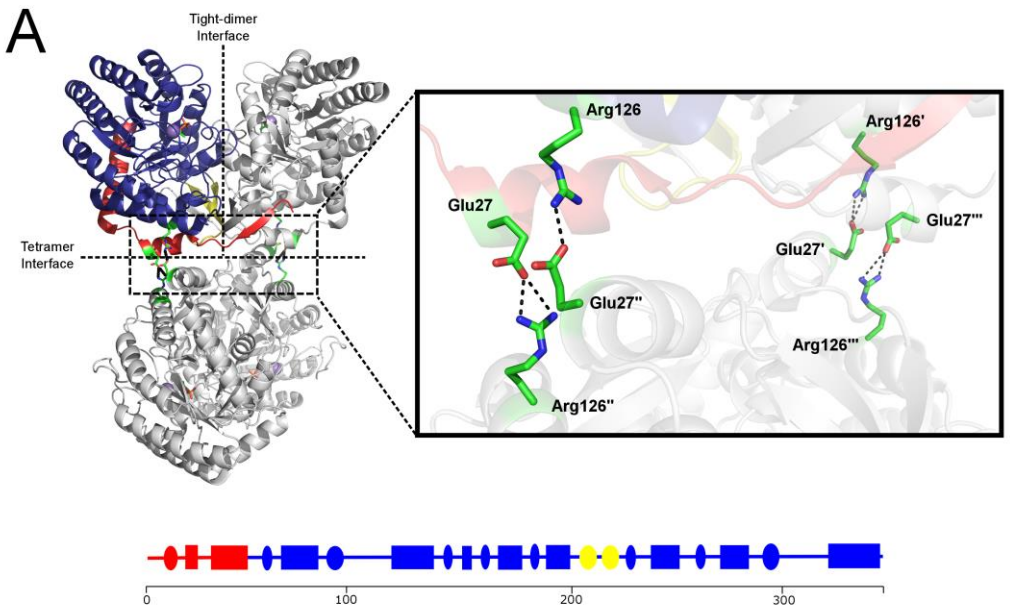
INTRODUCTION

The shikimate pathway is responsible for the biosynthesis of important aromatic compounds including the aromatic amino acids phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp).[18] The presence of this pathway in plants, microorganisms and apicomplexan parasites and its absence in mammals has drawn attention to the enzymes of this pathway as antimicrobial drug design targets [102]. 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) catalyses the first committed step of the shikimate pathway, which involves the aldol-like condensation of erythrose 4-phosphate (E4P) and phosphoenolpyruvate (PEP) to form 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAH7P). Entry into the pathway is controlled by the allosteric regulation of DAH7PS by the pathway end products Phe, Tyr and Trp, or other intermediates of the pathway .[20]

Despite significant sequence diversity, all known DAH7PS enzymes share a common $(\beta/\alpha)_8$-barrel core and similar active site architectures formed by conserved residues that support a similar catalytic mechanism.[35] On the basis of sequence and size, DAH7PS enzymes are classified into groups (type Iα, type Iβ and type II). Each group features distinct structural elements appended to the core catalytic barrel that are associated with the allosteric inhibition of the enzymes and different quaternary structure associations.[20,24,25,28,48,72,103] The type I groups, type Iα and type Iβ, share approximately 30% identity, and although a common dimeric unit is shared between by enzymes of these groups, they form different homotetrameric structures.[104]

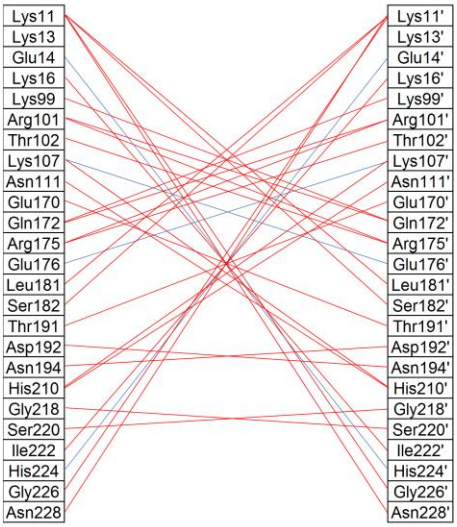
The DAH7PS isolated from *N. meningitidis (Nme*DAH7PS*)*, the causative agent of pyogenic meningitis and meningococcal septicaemia, belongs to the type Iα group. This protein adopts a homotetrameric assembly, and contains similar structural features to the other characterised type Iα DAH7PSs from *Saccharomyces cerevisiae* and *Escherichia coli*. *S. cerevisiae* and *E. coli* both express multiple isozymes of DAH7PS and the catalytic activity of each of these isozymes is sensitive to the presence of a different aromatic amino acid, however, the genome of *N. meningitidis,* encodes a single DAH7PS (*Nme*DAH7PS) which is most sensitive towards feedback inhibition by Phe.[26]

A comparison of the structure of the *S. cerevisiae* and *E. coli* enzymes with *Nme*DAH7PS reveals that although the single chain and dimeric units are superimposable, the two dimeric units that generate the homotetramer are oriented at different angles. The homotetrameric *Nme*DAH7PS comprises two interfaces, the more extensive (tight dimer) and the less extensive (tetramer interface) (Fig 1). The tight dimer interface involves interactions between 79 residues per chain and buries ~ 2800 $Å^2$ or 18 % of the surface area between chains. The tetramer interface is far weaker and consists of 17 residues per chain and buries only ~510 $Å^2$ or 3.8 % of the *Nme*DAH7PS chain. Both of these interfaces are comparable to the other structurally characterised type Iα DAH7PS enzymes.[22,33] In the *Eco*DAH7PS structure (PDB code: 1QR7) the angle between tight dimers is ~22˚ less than that of *Nme*DAH7PS (PDB code: 4HSN). Conversely, the arrangement of tight dimers in *Sce*DAH7PS (PDB code: 1OAB) is ~26˚ greater than that of *Nme*DAH7PS. This illustrates the variability at the tetramer interface between members of the same DAH7PS subfamily.



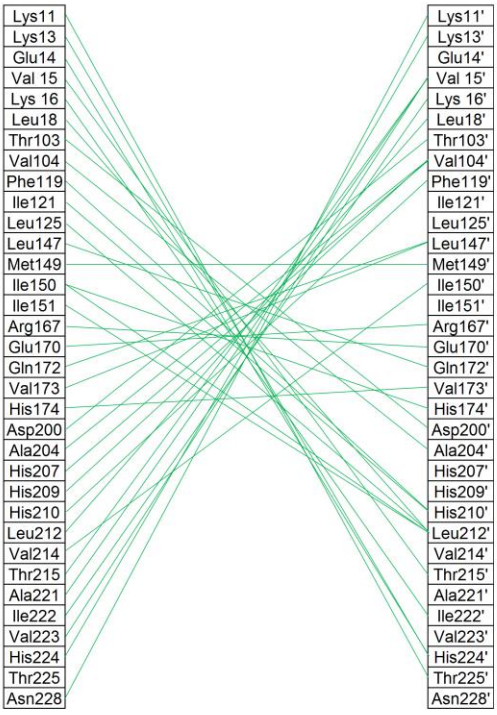
B

Nme DAH7PS tight dimer interface electrostatic interactions



C

Nme DAH7PS tight dimer interface hydrophobic interactions



Nme DAH7PS tetramer interface electrostatic interactions

Nme DAH7PS tetramer interface hydrophobic interactions

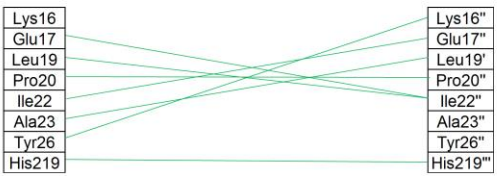


Fig 1. A). Quaternary structure of NmeDAH7PS (PDB code: 4HSN) indicating the tight dimer interface and tetramer interface. The core catalytic $(\beta/\alpha)_8$ barrel is coloured blue and the N-terminal and internal extensions are coloured red and yellow, respectively. Manganese metal ion (purple), ligand (PEP) and Glu27-Arg126 salt bridge (green-sticks) A schematic of secondary structural features of monomeric NmeDAH7PS according to residue number, circles denote β-sheets and rectangles denote α-helices B) Hydrogen bonding (red) and salt bridge (blue) interactions identified by analysis of each interface. Top. Tight-dimer. Bottom. Tetramer C) Hydrophobic interactions (green) at each interface of NmeDAH7PS. Top. Tight-dimer. Bottom. Tetramer . Chain B, C and D are denoted `, ``and ``` respectively.

Herein, we describe the disruption of the quaternary structure of *Nme*DAH7PS via mutagenesis of a single salt bridge at the tetramer interface, generating a dimeric species that we show to be the functional unit of *Nme*DAH7PS.

EXPERIMENTAL PROCEDURES

Bacterial strains, plasmids, media and growth conditions

Wild-type *Nme*DAH7PS

Purification procedures and kinetic measurements were performed as previously described.[26]

*Nme*DAH7PS^{R126S}

The variant R126S (*Nme*DAH7PS^{R126S}) was constructed using a Quik Change Lightning Mutagenesis Kit using the pT7-7 wild-type plasmid as template. Primers used to introduce the mutation were FOR: 5'-CATCAATTTCCGTTTGAGCCAGGAGAGGCAGCCTG-3' and REV: 5'-CAGGCTGCCTCTCCTGGCTCAAACGGAAATTGATG -3'. The PCR product was transformed into Top10 cells. Plasmids of the expected size, as determined by agarose gel, were DNA sequenced to identify those containing the successful mutation and transformed into BL21 Star^{TM} (DE3) cells. Expression and purification were carried out as previously described for the wild-type enzyme.[26]

Interface analysis

Resmap[105] and PISA[106] were used to analyse and visualise the interfacing residues. Free energy calculations for the formation of interfaces were calculated from PISA.

Differential Scanning Calorimetry

DSC was performed on a NanoDSC (TA instruments). Protein solutions were dialysed in SEC buffer (10 mM BTP, pH 7.3, 100mM KCl and 200µM PEP) and the dialysis buffer was used as reference. A final protein concentration of 26 µM (1 mg/mL) was used for the DSC experiments, and prior to scanning, all solutions were degassed by stirring under vacuum. An excess pressure of 3 atm was applied to the cells during scanning. Buffer scans were run prior to protein scans. DSC scans were performed using a heating rate of 1 °C/min, and the data was analysed using the Nanoanalyze™ software supplied with the instrument.

Analytical Size Exclusion Chromatography

To determine the molecular mass of *Nme*DAH7P synthase in solution, protein standards of known molecular mass were applied to a Superdex S200 column (GE Healthcare). β-Amylase (200 kDa), Alcohol dehydrogenase (150 kDa), ovalbumin (75 kDa), conalbumin (66 kDa) and carbonic anhydrase (29 kDa) were used as molecular weight standards.

Analytical ultracentrifugation

Sedimentation velocity experiments were performed in a Beckman Coulter XL-1 analytical ultracentrifuge equipped with UV/Vis scanning optics at 20 ^{o}C. Reference buffer solution (10 mM BTP, pH 7.3, 100 mM KCl, 200 μM PEP) and sample solutions were loaded into 12-mm double-sector cells with quartz windows, and mounted in an An-50 Ti 8-hole rotor.

Protein samples at concentrations of 0.2, 0.4 and 1.2 mg/mL were centrifuged at 50,000 *g* and the absorbance data was collected at 277 nm or 276 nm for *Nme*DAH7PS and *Nme*DAH7PS^{R126S} respectively. Radial scans collected between 5.8 cm and 7.2 cm with 0.003 cm increments were then analysed with the continuous size distribution model using SEDFIT.[107] The partial specific volume (*v*) of the protein samples (0.737734 mL/g), buffer density (1.005 g/mL) and buffer viscosity (0.01002 cp) were calculated using the program SEDNTERP.[108]

Small angle X-ray scattering

SAXS measurements – Measurements were performed at the Australian Synchrotron SAXS/WAXS beamline equipped with a Pilatus detector (1M, 170 mm x 170 mm, effective pixel size, 172 x 172 μm). The wavelength of the x-rays was 1.0332 Å. The sample–detector distance was 1600 mm, which provided an *q* range of 0.01–0.565 Å^{-1} [where *q* is the magnitude of the scattering vector, which is related to the scattering angle (2θ) and the wavelength (λ) as follows: $q = (4\pi/\lambda) \sin\theta$].

Data were collected from a 2 mm glass capillary (Hampton Research) and the temperature was 25 °C. Scattering data was collected from wild-type *Nme*DAH7PS and *Nme*DAH7PS^{R126S} (initial concentration of 7 and 12 mg/mL, respectively, prior to SEC) in the buffer (10 mM BTP, pH 7.5, 100 mM KCl, 200 μM PEP). Scattering from glucose isomerase at an initial concentration of 10 mg/ml was collected to calibrate the mass of scattering particles. Data were collected in 2 second intervals. 2D intensity plots from the peak of the SEC run were

radially averaged, normalised to sample transmission and background subtracted.

SAXS data analysis – Scattered intensity (*I*) was plotted *versus q*. Extrapolation of the DAH7PS *I*(*q*) profiles to zero angle [*I*(0)] and comparison with that of glucose isomerase standards indicated a molecular mass consistent with the *Nme*DAH7PS tetramer or dimer. All samples were devoid of an increase in intensity at low *q* (increase is indicative of aggregation). Radius of gyration (R_g) did not vary significantly over the concentrations measured, and all Guinier plots were linear for $q \cdot R_g < 1.3$. The data sets for structural analyses were recorded with 302 data points over the range $0.0110 \leq q \leq 0.4$ Å^{-1}. 1D profiles were background subtracted and Guinier fits were made using PRIMUS. Indirect Fourier transform was performed using GNOM[109] to yield the function *P*(*r*), which gives both the relative probabilities of distances between scattering centers and the maximum dimension of the scattering particle D_{max}. Theoretical scattering curves were generated from atomic coordinates and compared with experimental scattering curves using CRYSOL.[110]

Kinetics

The assay system for *Nme*DAH7PS was performed as previously described by monitoring the disappearance of PEP at 232 nm.[26] Standard assay buffer was 50 mM BTP (pH 6.8) containing $MnSO_4$ (100 µM) and variable concentrations of substrates PEP and E4P to a final volume of 1 mL. Assays were carried out at 25 °C.

*Nme*DAH7PS^{R126S} - The reaction mixtures for the determination of the K_m of E4P consisted of PEP (126 µM), $MnSO_4$ (100 µM) and E4P (9–140 µM), in 50 mM BTP, pH 6.8 buffer. The reaction mixtures for the determination of the K_m of PEP consisted of E4P (232 µM) and PEP (12–315 µM) in assay buffer. The reaction was initiated by the addition of *Nme*DAH7PS^{R126S} (2 µL, 1.3 mg/mL (final concentration-68 nM)).

*Nme*DAH7PS^{WT} - The reaction mixtures for the determination of the K_m of E4P consisted of PEP (100 µM) and E4P (10-320 µM), in assay buffer. The reaction mixtures for the determination of the K_m of PEP consisted of E4P (300 µM) and PEP (5-160 µM) in assay buffer. The reaction was initiated by the addition of *Nme*DAH7PS^{WT} (2 µL, 0.6 mg/mL (final concentration-32 nM)). Apparent K_m and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Grafit (Erithicus).

Time dependent loss of activity

Standard assays were performed which consisted of E4P (230 µM) and PEP (100 µM) in assay buffer to a final volume of 1 mL. Assays were initiated by the addition of 2µl of ~1 mg/mL *Nme*DAH7PS^{WT} or *Nme*DAH7PS^{R126S} (final concentration-52 nM). Assays were performed in triplicate and averaged at timepoints (0 hr, 2 hr, 4 hr, 6 hr, 18 hr and 24 hr). Both *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S} were diluted to 1 mg/mL at time 0 hr and kept at ~4 °C for the experiment duration. There was no loss of protein after 24 hr at 1 mg/mL calculated via absorbance at 280 nm using the calculated molar extinction coefficient of 31650 M^{-1} cm^{-1} post experiment.[111]

Inhibition

Standard assays were performed which consisted of E4P (74 µM), PEP (162 µM) and the amino acid; Phe, Tyr or Trp (25 µM – 1 mM) in assay buffer. Assays were were initiated by the addition of *Nme*DAH7PS^{R126S}. Assays were performed in triplicate.

Isothermal titration calorimetry

Binding of *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S} to Phe at pH 7 was measured by ITC using a VP- ITC unit operating at 298 K (Microcal, GE Health- care). Before use, the protein was buffer exchanged against binding buffer [0.5 mM $MnSO_4$ in 50 mM BTP buffer (pH 7)] and all solutions were degassed in a vacuum. Protein concentration was measured by UV absorption immediately before titrations were started. The titrations were comprised of 28 injections, one 2 µL injection followed by 27 x 10 µL injections of Phe. The initial datum point was routinely deleted to allow for diffusion of ligand across the needle tip during the equilibration period. A heat of dilution experiment was measured independently and subtracted from the integrated data before curve fitting in Origin 7.0. For the binding of *Nme*DAH7PS to Phe, 15 µM of *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S} were used and the syringe contained 1.2 mM Phe; and the data were fitted with the two-site sequential-binding model supplied by Micro-Cal.

Crystallisation of *Nme*DAH7PS^{R126S}

*Nme*DAH7PS^{R126S} was expressed and purified as previously described for the wild-type enzyme.[26] The resulting protein solution was concentrated to approximately 10 mg/ mL. 1 µL

of enzyme solution (9–11 mg/mL) was mixed with 1 µL of crystallisation buffer containing 0.1 M Tris HCl (pH 7.3), 0.2 M trimethyl-amino-*N*-oxide (TMAO), 600 µM $MnSO_4$ and 15–20 % (w/v) PEG 2000MME. Crystals were grown by hanging drop vapour diffusion over 500 µL of crystallisation buffer and the crystallisation trays were incubated at 20 ºC. Crystals began to form in 48 hours and were fully formed within seven days. Crystals were flash frozen using liquid nitrogen in a cryoprotectant solution containing reservoir solution and 20 % (v/v) PEG400.

Crystallography and structure determination

An X-ray diffraction dataset was collected at the Australian Synchrotron using the MX1 beamline.[112] The datasets were integrated and processed using XDS and Aimless.[113,114] Appropriate cut-off resolution was determined via $CC^{1/2}$≥0.5 ensuring the data was complete in the highest resolution shell.[115] Space group and unit cell parameters for *Nme*DAH7PS^{R126S} were the same as those previously identified for the wild-type enzyme (PDB code 4HSN[26]) indicating that initial phases could be obtained via molecular replacement using the original structure as a search model in Phaser MR.[116] All ligands and waters were removed from the search model (PDB code 4HSN) before molecular replacement was carried out. Refmac5 was used to generate the electron density map and this was manually analysed and refined in COOT (Table 1).[117–119] The quality of the model was optimised by consecutive model building in COOT and refinement with Refmac5. Water molecules were added manually via interpretation of the |2*Fo-Fc*|map ensuring that they had the ability to hydrogen bond to at least one acceptor or donor. In all structures, no electron density for the 14 amino acids at the N-terminus of each *Nme*DAH7PS^{R126S} chain could be found, which is attributed to these residues being part of a highly flexible allosteric region of the protein. Molprobity was used to assess structure quality during refinement cycles and before deposition.[120]

Table 1. Data collection and refinement statistics of NmeDAH7PS^{R126S} variant (PDB code: 4UCG)

	***Nme*DAH7PS^{R126S}**
Data Collection	
Crystal system; space group	Orthorhombic, $P2_12_12_1$
Unit cell parameters	
a, b, c (Å)	78.95, 132.7, 147.7
α, β, γ (º)	90, 90, 90
Resolution range (Å)	48.03-2.00 (2.03-2.00)
Measurements	1558648
Unique reflections	105375
Redundancy	14.8
Completeness (%)	100.0 (99.9)
I/σ (I)	15.2 (1.5)
R_{merge}	0.152
$CC_{1/2}$	0.564
Wilson *B*-value (Å)^2	28.94
Matthews coefficient	2.54
Refinement	
R_{factor}	0.1801
R_{free}	0.2200
Chain length	351
Observed number of residues	334 (Chain B & D), 335 (Chain C), 336 (Chain A)
Water molecules	821
Other (PEG's, Mn, $SO4^-$)	10
Ligand	4
Mean *B* (Å)^2	
Protein	35.02
Water	40.5
Other	55.6
Ligand (PEP)	29.63 at 0.75 occupancy
r.m.s.d from target values	
Bond lengths (Å)	0.0100
Bond angles	1.3466
Dihedral angles	0.0749
Ramachandran	
Preferred (%)	97.55
Allowed (%)	1.84
Outliers (%)	0.61
PDB Entry	4UCG

RESULTS

Interface analysis of interface and choice of mutation site

The minor interface through which the tetramer is formed comprises primarily hydrophilic interactions, in contrast to the extensive hydrophobic interactions that contribute to the tight dimer interface (Fig 1). Of note is the single salt bridge between Arg126 and Glu27 from opposing tight dimers, which may play a role in determining the twist between tight dimers (Fig 1). *E. coli* DAH7PS shares a similar tetramer interface to that of *Nme*DAH7PS, and contains the same Glu-Arg salt-bridge. This salt bridge is not observed for *S. cerevisiae D*AH7PS, whose tetramer interface consists solely of hydrophobic interactions between four residues from each chain. This salt bridge was disrupted by substitution of Arg126 to a serine residue and this change was chosen to maintain the hydrophilicity of the protein surface.

The variant protein *Nme*DAH7PS^{R126S} was created using site directed mutagenesis and purified following the protocol developed for the wild-type protein (*Nme*DAH7PS^{WT}). *Nme*DAH7PS^{R126S} had an expected mass of 77,300 Da, and showed similar thermal stability properties to the wild-type protein, with a melting temperature measured by differential scanning calorimetery of 47 ˚C (*Nme*DAH7PS^{WT} displayed a melting temperature of 49˚C) (S1 Fig).

*Nme*DAH7PS^{R126S} is dimeric in solution

Analysis of the quaternary structure assembly of the *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S} enzymes was carried out using both analytical size-exclusion chromatography (SEC) and analytical ultracentrifugation (AUC). SEC analysis revealed a clear difference in the solution behaviour of the two enzymes (Fig 2A). The wild-type protein gave a clear symmetrical peak shape corresponding to a mass of 150 kDa. Based on the theoretical molecular weight of a single wild type enzyme chain (38.7 kDa), this observation is consistent with a homotetrameric protein. Conversely, for *Nme*DAH7PS^{R126S} variant no tetrameric species was observed. An asymmetrical peak corresponding to a mass of ~66 kDa is consistent with the presence of a dimer under the conditions used for analysis. The same profile for each enzyme was observed when the samples were prepared for analysis and left at room temperature for 18 hours, indicating that the two enzymes observed form stable quaternary forms.

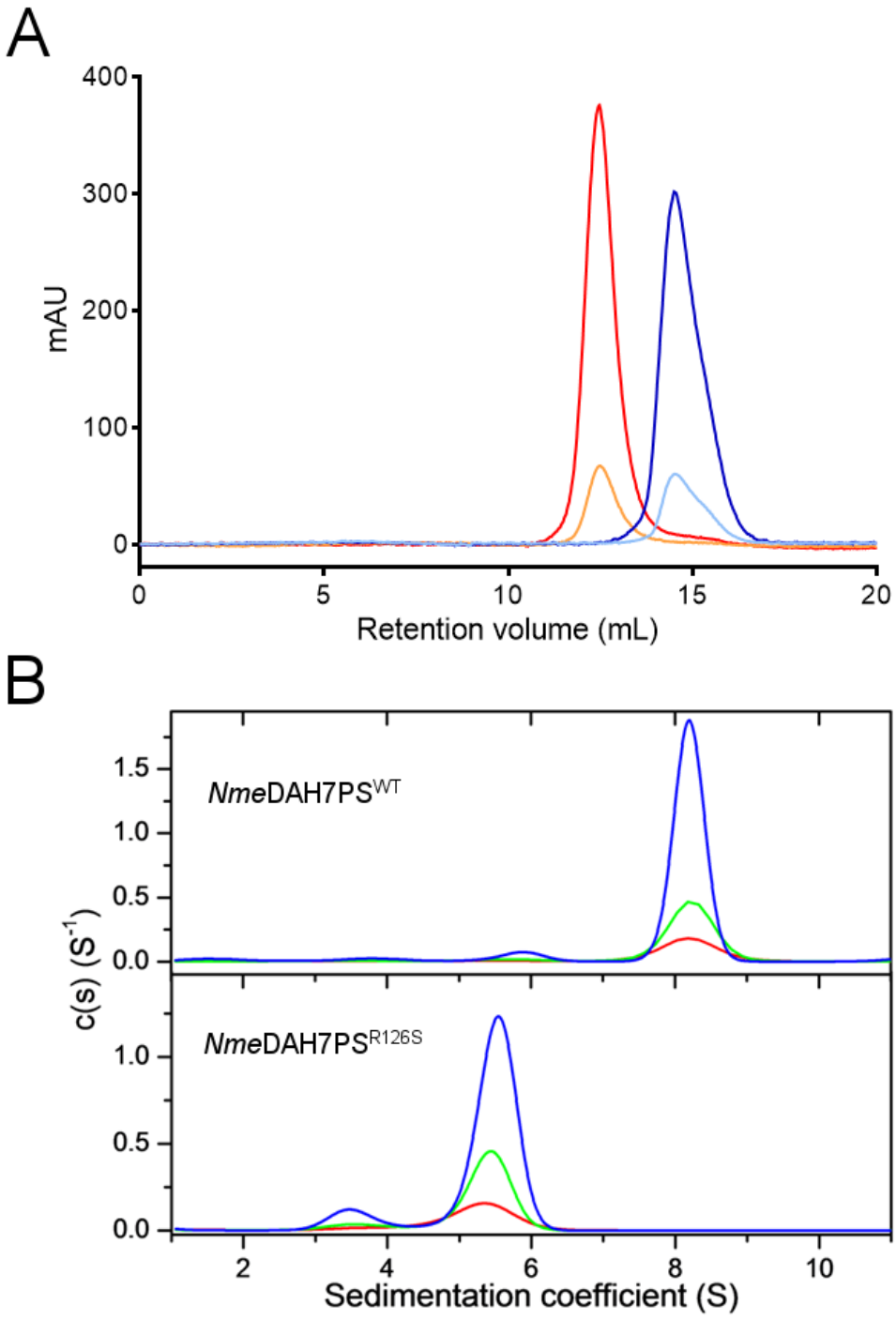


Fig 2. (A) Analytical SEC traces of NmeDAH7PS^{WT} and NmeDAH7PS^{R126S} at two different concentrations. NmeDAH7PS^{WT} 1 mg/mL (red) and 0.2 mg/mL (orange). NmeDAH7PS^{R126S} 1 mg/mL (blue) and 0.2 mg/mL (light blue). B. Calculated sedimentation coefficient c(s) distribution plot of wild-type NmeDAH7PS (top) and NmeDAH7PS^{R126S} (bottom). For both enzymes, the distribution of species for three different concentrations are shown: 0.2 mg/mL (red), 0.4 mg/mL (green) and 1.2 mg/mL (blue). The sedimentation velocity data, fit data and residuals from the fits are shown in the supplementary material (S2 Fig).

Sedimentation velocity experiments were carried out on *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S} enzymes. The absorbance versus radial position data for the *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S} variant are shown in Fig 2B. Fitting the data for *Nme*DAH7PS^{WT} to a continuous size distribution model indicates the presence of a major species with a standardised sedimentation coefficient of 8.1 S (Fig 2B). The calculated mass of this species is 141 kDa, which is consistent with the expected mass of a homotetrameric species (theoretical mass of 154.6 kDa). For *Nme*DAH7PS^{R126S}, the fit indicated the presence of a major species with a standardised sedimentation coefficient of 5.7 S (Fig 2B). This predominant species has a calculated mass of 69 kDa, consistent with the expected molecular weight of the homodimeric assembly (77.3 kDa). At all three concentrations measured (0.4, 0.8, 1.2 mg/mL), a minor species with a sedimentation coefficient of 3.7 S, corresponding to an expected molecular weight of 36 kDa was observed, which is consistent with the molecular weight of monomeric *Nme*DAH7PS (38.7 kDa).

Small-angle X-ray scattering (SAXS) was employed to confirm the quaternary structure of *Nme*DAH7PS R126S in solution. The scattering profile collected for the *Nme*DAH7PS^{R126S} variant was compared to that collected for the wild-type enzyme (Fig 3, Table 2)). The difference between the scattering intensities at low angles is indicative of a change in molecular weight and suggests these two enzymes adopt distinct oligomeric assemblies. The radius of gyration (R_g) calculated from the scattering pattern of *Nme*DAH7PS^{R126S} variant using the indirect Fourier transform function, ($P(r)$), was ~ 26 Å and the maximum diameter of the scattering particle, D_{max}, was calculated as ~ 95 Å using GNOM.[109] These values are in close agreement with the tight-dimer generated from the *Nme*DAH7PS^{R126S} crystal structure (R_g of ~ 25 Å and D_{max} of ~ 86 Å). Similarly, CRYSOL[110] was used to directly compare the collected data to the hypothetical scattering from the *Nme*DAH7PS^{R126S} tetramer and dimer generated from the crystal structure. These data clearly show the *Nme*DAH7PS^{R126S} variant adopts a dimeric structure in solution comparable to that of the *Nme*DAH7PS^{R126S} tight-dimer (Fig 3B).

Table 2: SAXS parameters for NmeDAH7PS^{WT} and NmeDAH7PS^{R126S} variant.

Structural parameters	***Nme*DAH7PS^{WT}**	***Nme*DAH7PS^{R126S}**
R_g (Å)[from $P(r)$]	38 ± 1	27 ± 1
R_g (Å)[Guinier]	37 ± 1	27 ± 1
D_{max} (Å)[from $P(r)$]	110 ± 5	95 ± 5
$I(0)$ (cm^{-1})[from Guinier]	0.113 ± 0.01	0.090 ± 0.01
Hydrated volume (Da^3) [from *P(r)*]	224,000 ± 230	110,900 ± 110
Theoretical volume (Da^3)[from CRYSOL]	222,900	101,300

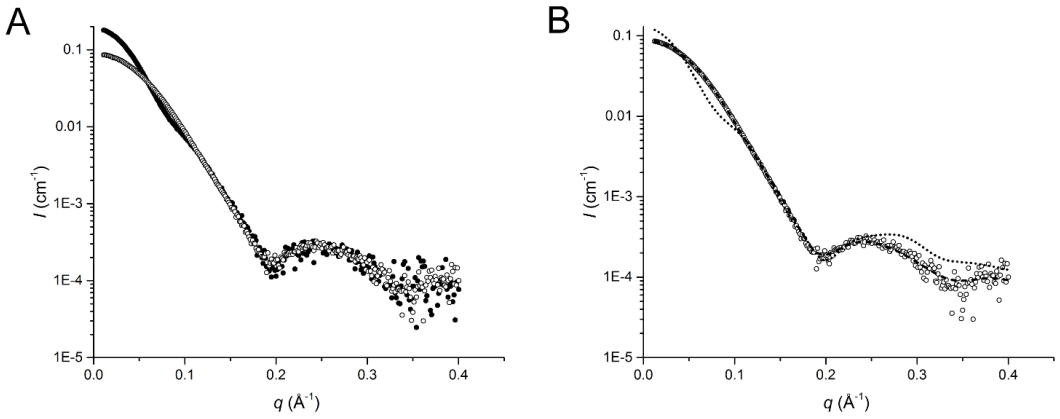


Fig 3. SAXS scattering profiles collected for NmeDAH7PS^{WT} and NmeDAH7PS^{R126S}. (A) Overlay of the scattering data of NmeDAH7PS^{WT} (closed circles) and NmeDAH7PS^{R126S} variant (open circles) and (B) CRYSOL fits of the NmeDAH7PS^{R126S} variant data to the theoretical scattering data generated from the R126S tetramer (dotted line) and tight-dimer (dashed line). The discrepancies of the fit of the theoretical scattering of tetrameric and dimeric forms to the experimentally determined data (χ) were calculated as 22.80 and 0.67 respectively.

*Nme*DAH7PS^{R126S} is active

*Nme*DAH7PS^{R126S} was observed to be catalytically active with parameters that are comparable with those of the wild-type enzyme (Table 3). The apparent K_m^{PEP} of *Nme*DAH7PS^{R126S} is 25 µM, which is similar to that observed for the wild-type enzyme (15 µM). The *Nme*DAH7PS^{R126S} displays a lower K_m^{E4P} (13.4 µM) than the wild-type enzyme (37 µM). However, there is a noticeable decrease in specific activity for the dimeric variant. This implies that tetramerisation may have some influence on catalysis by *Nme*DAH7PS.

Table 3. Kinetic parameters of NmeDAH7PS WT and NmeDAH7PS^{R126S}.

Organism	$K_{m(PEP)}$ (μM)	$K_{m(E4P)}$ (μM)	k_{cat} (s^{-1})	$k_{cat}/K_{m(PEP)}$	$k_{cat}/K_{m(E4P)}$
***Nme*DAH7PS^{WT}**	15 ± 1	37 ± 2	27.1 ± 0.1	1.8	0.7
***Nme*DAH7PS^{R126S}**	25 ± 1	13.4 ± 0.3	12.6 ± 0.1	0.5	0.9

Apparent K_m values were determined at high concentrations of the second substrate.

*Nme*DAH7PS^{R126S} shows time-dependent loss of activity

Comparison of stability of *Nme*DAH7PS^{WT} and the *Nme*DAH7PS^{R126S} was carried out by measuring the activity of the enzymes over a 24 hour period. *Nme*DAH7PS^{WT} lost ~10% catalytic activity, whereas the dimeric variant was far less stable and was observed to retain less than 60% of its activity after 24 hours (Fig 4).

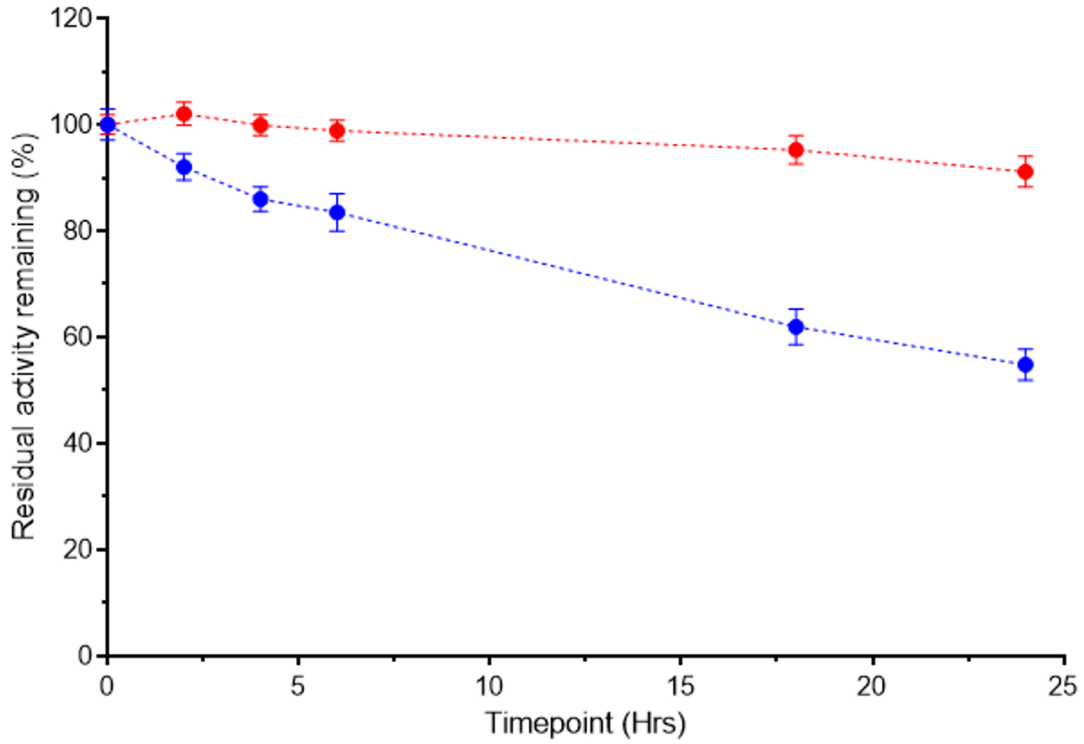


Fig 4. Catalytic activity of NmeDAH7PS^{R126S} (blue) NmeDAH7PS^{WT} (red) over 24 hour time period relative to maximum activity observed for each enzyme at the beginning of the experiment.

*Nme*DAH7PS^{R126S} binds and is inhibited by Phe

As for the wild-type enzyme, *Nme*DAH7PS^{R126S} is inhibited to the largest extent by Phe and to a lesser extent by Tyr and Trp (Fig 5A). Isothermal titration calorimetry experiments were employed to illustrate that the affinity of the dimeric enzyme for Phe, was not altered upon mutation (Fig 5B). For *Nme*DAH7PS^{R126S} cooperativity was observed in the binding isotherm.

The data were fitted to a two-site sequential model giving K_D values for *Nme*DAH7PS^{R216S} of 1.2 ± 0.7 µM for the first binding event and 33 ± 3 µM for the second binding event. This indicates similar Phe binding characteristics to those observed for *Nme*DAH7PS^{WT} [26].

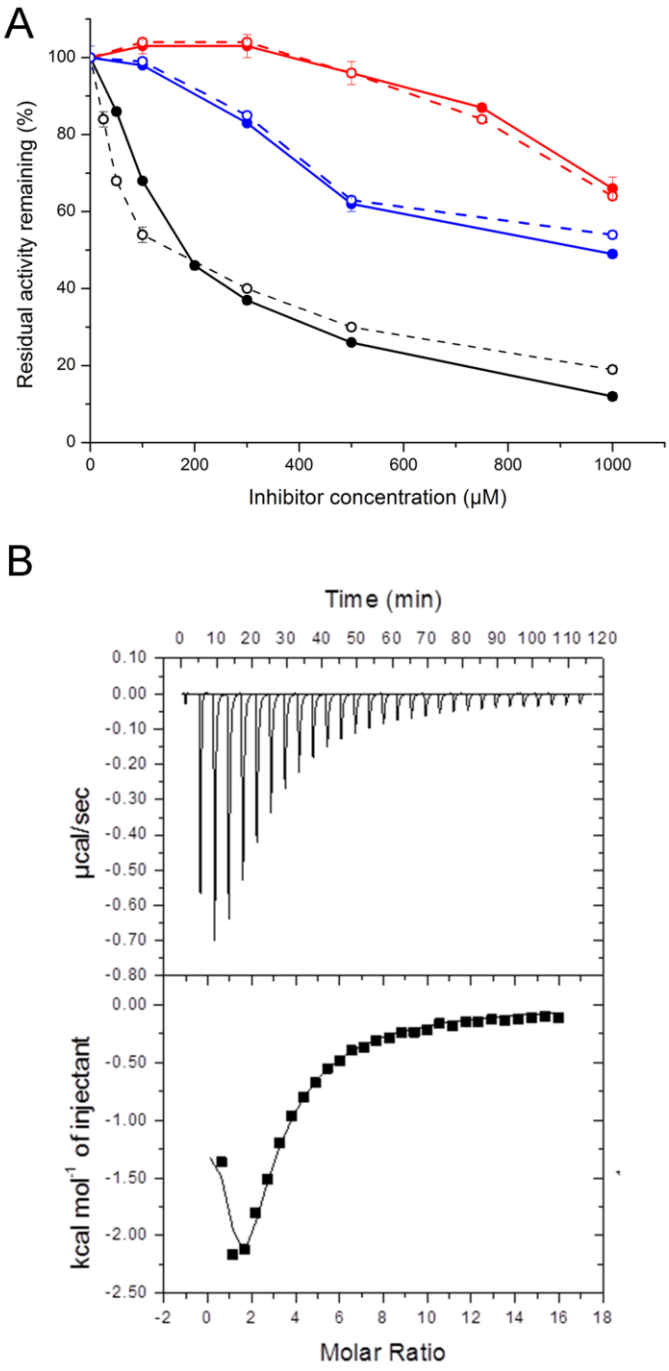


Fig 5. A. Effect of the aromatic amino acids, Phe (black), Tyr (red) and Trp (blue) on the catalytic activity of NmeDAH7PS^{R126S} (open circles) in comparison to data previously reported for the wild-type enzyme (closed circles).[26] B. Isothermal titration calorimetry binding isotherm for NmeDAH7PS^{R126S} fit to a two-site sequential binding model.

*Nme*DAH7PS^{R126S} crystallises as a tetramer

The crystal structure of the *Nme*DAH7PS^{R126S} variant was determined by X-ray diffraction. This structure was refined to a resolution of 2.0 Å and crystallised in an orthorhombic space group ($P2_12_12_1$). Data collection and refinement statistics are shown in table 1. This space group is different to *Nme*DAH7PS^{WT} (PDB code 4HSN), which crystallised in the monoclinic ($P12_11$) space group.[26]

The asymmetric unit for *Nme*DAH7PS^{R126S} structure (PDB code: 4UCG) contains four chains interacting to provide a homotetramer, resembling the arrangement of wild-type *Nme*DAH7PS. Like the wild-type enzyme, the *Nme*DAH7PS^{R126S} tetramer has two distinct interfaces between the chains. The tight dimer interface remains extensive, burying 14% of the surface area of the monomeric unit, but is smaller than *Nme*DAH7PS^{WT} (18%). The tetramer interface, on which the R126S substitution is located, is far smaller, burying 3.8% of the surface of the monomeric unit in *Nme*DAH7PS^{WT} and this is reduced to 3.4% for the *Nme*DAH7PS^{R126S}. Superposition of the *Nme*DAH7PS^{R126S} homotetramer on the *Nme*DAH7PS^{WT} protein yields an RMSD of 0.61 Å (1328 Cα atoms), whereas superposition of *Nme*DAH7PS^{R126S} dimeric units on those of the wild-type protein gives RMSD values between 0.18-0.25 Å. These values are consistent with the small decrease in rotation of ~2˚of the dimeric units with respect to each other at the tetramer interface, and this change may also account for the difference in space group between the *Nme*DAH7PS^{R126S} and *Nme*DAH7PS^{WT} structures.

According to PISA server calculations[106], the *Nme*DAH7PS^{R126S} crystal structure displays a theoretical solvation free energy gain (Δ*G*) for the formation of the tight dimer interface of -16.1 and -16.5 kcal/mol, correlating to the chain A-chain B and chain B-chain D interfaces respectively[106]. This value is less that what is observed for the *Nme*DAH7PS^{WT} enzyme, where the Δ*G* values are -20.1 kcal/mol and -18.1 kcal/mol. The lower values are indicative of a weakening of the tight dimer interface in *Nme*DAH7PS^{R126S}.

The Δ*G* for formation of the tetramer interface in *Nme*DAH7PS^{R126S} are -6.5 and -9.1 kcal/mol for the chain A-chain C and chain B-chain D interfaces respectively (interface interactions are shown in S3 Fig, supplementary material). Comparison of this with the Δ*G* for formation of the

tetramer interface in *Nme*DAH7PS^{WT} of -6.9 and -7.0 kcal.mol, provides interesting insight into the structural asymmetry at the tetramer interface in *Nme*DAH7PS^{R126S}. The structural asymmetry observed in the crystal structure of *Nme*DAH7PS^{R126S} may contribute to the in-solution dimerisation of this variant. This directly implicates the Arg126-Glu27 salt bridge in the formation of a symmetrical tetramer interface in *Nme*DAH7PS^{WT}. Comparison of the surrounding residues in the crystal structure of *Nme*DAH7PS^{R126S} with that of *Nme*DAH7PS^{WT} illustrates this structural asymmetry. Tyr26, adjacent to Glu27, adopts a new, or alternate conformation in three of the four chains and in this new conformation, it forms a hydrogen bond with the backbone amine of Glu17. This position was previously occupied by a water molecule in the *Nme*DAH7PS^{WT} structure (Fig 6). Furthermore, His219, located at the centre of the tetramer, adopts several alternative conformations in the *Nme*DAH7PS^{R126S} structure as a result of this asymmetry.

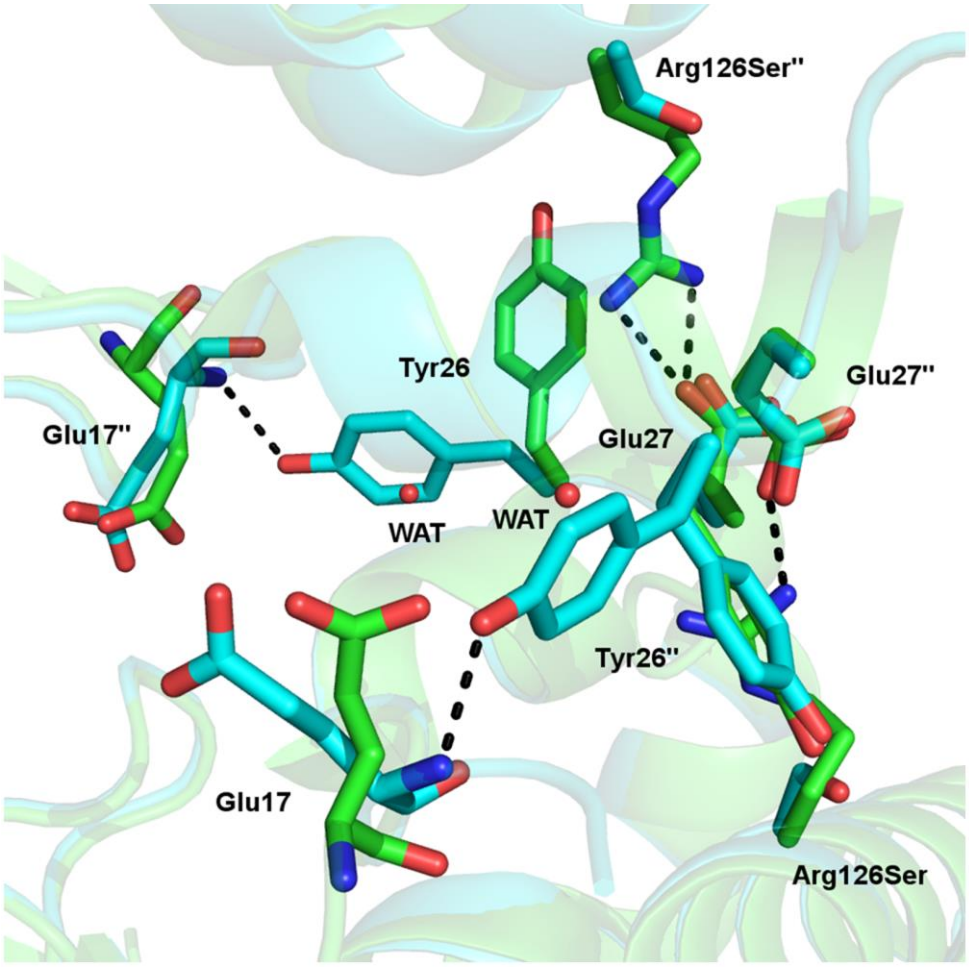


Fig 6. Tetramer interface of NmeDAH7PS^{WT} (green) (PDB code:4HSN) and NmeDAH7PS^{R126S} (cyan) indicating the variable positioning of Tyr26 in place of the conserved water molecule from the NmeDAH7PS^{WT}. The loss of salt bridge between Glu27 and Arg126 for the NmeDAH7PS^{R126S} variant is also illustrated.

DISCUSSION

A key feature of proteins is their tendency to form multimeric structures, usually homomers with even numbers of subunits all derived from the same gene, generating symmetry related interfaces, resulting in a functional biological assembly.[75] The DAH7PS family of proteins adopts a variety of homotetrameric forms, and like many other homotetramers, these tetramers are dihedral in nature and comprise dimers of dimers. Whereas the type Iα proteins share a common dimer subunit with the type Iβ group, they have a unique, small tetramer interface.

A single residue exchange in the small tetramer interface resulted in a solution- state dimer. The presence of this dimer and the absence of tetramer were confirmed by three independent solution-state techniques. The SAXS measurements also confirm that, as expected, the dimeric form adopts the assembly of the dimeric assembly of the parent protein. In contrast to these-solution state measurements, the protein crystallises as a tetramer. Some minor disruption and asymmetry has been introduced by the substitution indicating that this interface may be less stable. It should be noted that the dimeric form was observed by SAXS measurements even at relatively high concentrations, suggesting the crystal packing and crystallisation conditions favour the tetrameric form. Similar observations have been made with other related systems. The equivalent salt bridge of the *E. coli* Phe sensitive enzyme was disrupted by the mutation of Glu24 to Gln. This also resulted in a tetrameric crystalline form, however the enzyme was reported to be dimeric in solution, although no functional studies of this protein were provided.[31] Likewise, the tetrameric type Iβ enzyme from *Pyrococcus furiosus* has also been disrupted by mutation to form a mixture of dimeric and tetrameric forms in solution.[104] In this case the protein also crystallised as the tetramer.
The dimeric *Nme*DAH7PS retained its catalytic and allosteric properties. Catalysis was observed and both Phe binding and response was measured, revealing that the primary functional unit can be considered to be dimeric. The tetrameric form appears to provide advantage by a boost in catalytic rate and increased protein stability; the *Nme*DAH7PS^{R126S} enzyme lost significantly more catalytic activity over a long period than the wild-type tetrameric protein, and this instability may contribute to the reduction in k_{cat}. Once again parallels can be drawn here from the studies with the *P. furiosus* DAH7PS. The thermal

stability of this protein, derived from a hyperthermophilic source, was significantly reduced by disruption of the tetramer .[104]

These studies may help shed light on the common evolution and evolutionary divergence of the DAH7PS proteins into the different classes. There is some evidence that dihedral homotetramers become established through the interaction of symmetrical dimers for which the largest interface is included. It has been shown that the symmetrical nature of the tetramer interface in *Nme*DAH7PS may be a contributing factor to tetramerisation. It is becoming increasingly clear is that both type Iα and type Iβ DAH7PS proteins are functional through a similar dimeric structure. Type Iα proteins differ in the extent of this interface, as the allosteric machinery including an N-terminal extension and a β-hairpin contributes to the interactions that form the extensive dimer interface and provide the binding site for Phe. Furthermore, it is notable that the binding of Phe to both the tetrameric and dimeric forms of *Nme*DAH7PS is inhibitory and follows a two-site sequential model suggesting that both the allosteric and catalytic function can be provided by the dimeric unit.

This investigation clearly illustrates the importance of a single salt-bridge at tetramer interface of *Nme*DAH7PS and has implications for the evolution of quaternary structure in this family of enzymes.

Acknowledgements

Part of this research was undertaken on the MX and SAXS/WAXS beamlines at the Australian Synchrotron, Victoria, Australia with access administered by the New Zealand Synchrotron Group. L.C.H is grateful for a BIC administered Marsden funded PhD Doctoral scholarship. This research was funded, in part, by the New Zealand Marsden Fund (UOC1105).

Supplementary material

S1 Fig. Differential scanning calorimetry showing thermal stability of *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S}.

S2 Fig. Sedimentation velocity data for *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S}.

S3 Fig. Interface analysis of *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S}.

CHAPTER 4 PART II-The role of the central histidines

4.2 INTRODUCTION

Of the twenty natural amino acids, histidine (His) is the most versatile. It can form π-π stacking, cation-π and hydrogen-π interactions with His as well as other aromatic amino acids. Furthermore, it also plays a major role in the coordination of metal ions. In proteins, the pK_a of His has a very broad range, meaning that it can play the role of both donor and acceptor.[121]

The Arg126-Glu27 salt bridge across the Chain A-C and Chain B-D interface has been shown to be critical to the solution state tetrameric structure of *Nme*DAH7PS. However, there is one interaction across the diagonal interface of the *Nme*DAH7PS (Figure 4.2.1). The area of contact between chains A and D (B and C) is limited to ~30 Å^2 (~0.2% of solvent-accessible area of the single chain). This consists of a π-stacking interaction at a distance of ~3.6 Å between the imidazole rings of His219, located on the loop connecting the β_{6a} and β_{6b} strands of the β-hairpin insertion. The four His219 residues at the intersections of the two-fold axis are intriguing and contribute to the interaction between the two tight dimers. They may also have an important role in the twist of the tight dimers relative to one another at the tetramer interface. In addition, there is also a shared proton forming a hydrogen bond between the Nε atoms of the His219 residues of chains A and C (B and D). The four histidines that make up this interaction in *Nme*DAH7PS are conserved for the Phe sensitive isozyme of *Eco*DAH7PS and show very similar characteristics.[50] The location of these histidines has recently been hypothesised as a possible novel allosteric site for *Eco*DAH7PS.[122]

However, in the presence of allosteric or active site inhibitor, one or both of these π-stacking interactions are lost. In the crystal structures containing the inhibited form of DAH7PS (PDB code: 4UC5, 4UMA, 4UMB, 4UMC) the central histidine residues (Figure 4.2.1) adopt several conformations implying an increase in the flexibility of these residues and a change in the association between the tight dimers at the tetramer interface.[100]

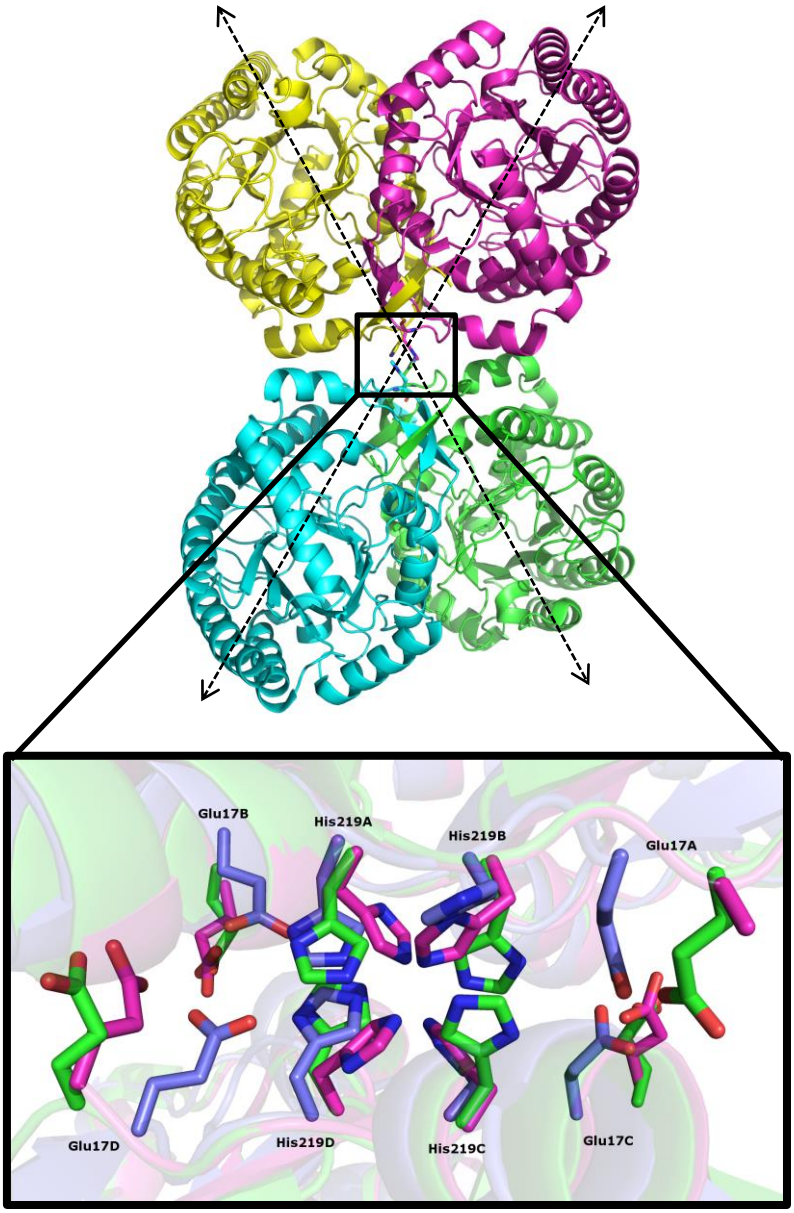


Figure 4.2.1 NmeDAH7PS^{WT} tetramer indicating the superposition of the E-vinyl phosphonate bound (PDB code: 4UMA) (magenta), Phe-bound (PDB code: 4UC5) and PEP-bound (PDB code: 4HSN) (green) illustrating the multiple rotamer positions that His219 adopts when an inhibitor is bound at either site. The two-fold axes interaction is depicted with double headed arrows.

There is unexplained electron density in the $|F_o\text{-}F_c|$ maps for the region including these residues when the enzyme is in its inhibited state (PDB code: 4UC5 and 4UMA). This is in accordance with the ability of a molecule, ligand or metal ion to bind here. In the Phe-bound crystal structure a single sulphate molecule has been modelled into the $|F_o\text{-}F_c|$ maps, however, when the enzyme has an active-site inhibitor bound such as the *E*-vinyl phosphonate, the residues either coordinate a molecule, metal ion, ligand or adopt several conformations (Figure 4.2.2).

It has been shown that upon disruption of the tetramer, *Nme*DAH7PS does not experience any major functional changes. However, the calculated theoretical solvation free energy gain for the formation of the tight dimer interface was reduced in the dimeric variant. Mutation of the only interaction across the diagonal interface of the tetramer may illustrate the importance of tetramer formation in *Nme*DAH7PS and the importance of the dihedral angle between tight dimers. This alteration may have an effect on the ability of the allosteric binding loop to close as there is a hydrogen bond formed between Glu17, the backbone carbonyl of Leu18 and the sidechain N of His219 in the Phe-bound structure (PDB code: 4UC5). It may also have a direct effect on the interactions at the interface as the allosteric modifications of *Nme*DAH7PS (N-terminal extension-Glu17, Leu18 and the β-hairpin insertion-His219) are crucial for the formation of both the tight dimer and tetramer interfaces. The interaction of the β-hairpin insertion with the catalytic $\beta_2\alpha_2$ loop observed in the PEP-bound *Nme*DAH7PS structure (PDB code: 4HSN) may implicate this residue (also present in the Phe sensitive isozyme from *Eco*DAH7PS) in catalysis as well as the conferral of regulation in *Nme*DAH7PS.

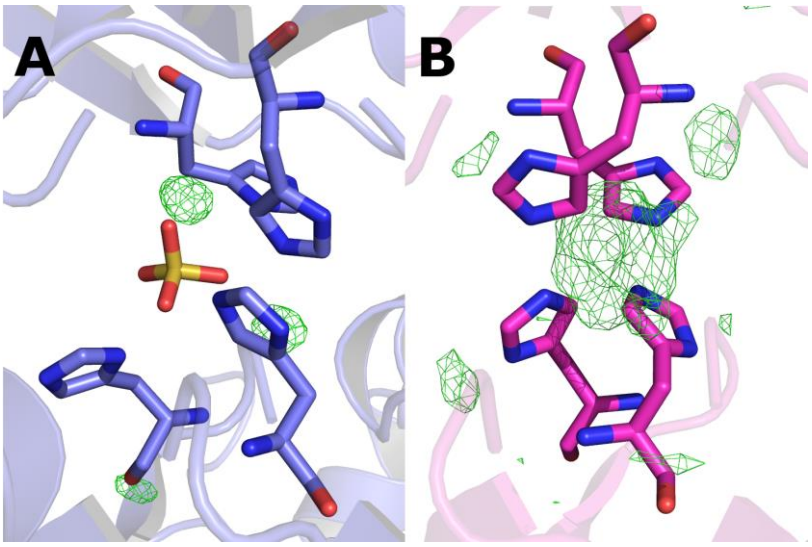


Figure 4.2.2 $|F_o-F_c|$ maps of the four His219 residues. A) Phe-bound crystal structure (PDB code: 4UC5) indicating the modelling of a sulphate molecule into the position between the flexible residues. B) E-vinyl phosphonate crystal structure (PDB code: 4UMA) indicating the unmodelled $|F_o-F_c|$ electron density at the centre of the histidine residues and the coordinating geometry undertaken by each His219

4.3 RESULTS

4.3.1 Kinetics and inhibition of *Nme*DAH7PS^{His219Ala} variant

The kinetic parameters of the *Nme*DAH7PS^{His219Ala} variant enzymes were determined under the conditions stated for *Nme*DAH7PS^{WT} (section 7.1.27).

	PEP K_M (µM)	E4P K_M (µM)	k_{cat} (s^{-1})	$k_{cat}/K_{M\ (PEP)}$	$k_{cat}/K_{M\ (E4P)}$
***Nme*DAH7PS^{His219Ala}**	17 ± 1	24 ± 2	8.9 ± 0.1	0.5	0.4
***Nme*DAH7PS^{WT}**	15 ± 1	37 ± 2	27.1 ± 0.2	1.8	0.7

Table 4.3.1.1 Kinetic parameters of NmeDAH7PS^{His219Ala} variant enzyme

The kinetic characterisation of *Nme*DAH7PS^{His219Ala} indicates that the mutation does not have a large effect on the kinetic parameters for this enzyme. The K_M for PEP remains constant and the K_M for E4P is slightly less than that of the *Nme*DAH7PS^{WT}. However, the k_{cat} is reduced compared to *Nme*DAH7PS^{WT} (Table 4.3.1.1).

The inhibitory effect of Phe on the *Nme*DAH7PS^{His219Ala} variant is very similar to that observed for the *Nme*DAH7PS^{WT} enzyme at all concentrations of Phe tested (Figure 4.3.1.1). This result implies that the communication of allosteric signal to the active site is not mediated in any way by this residue.

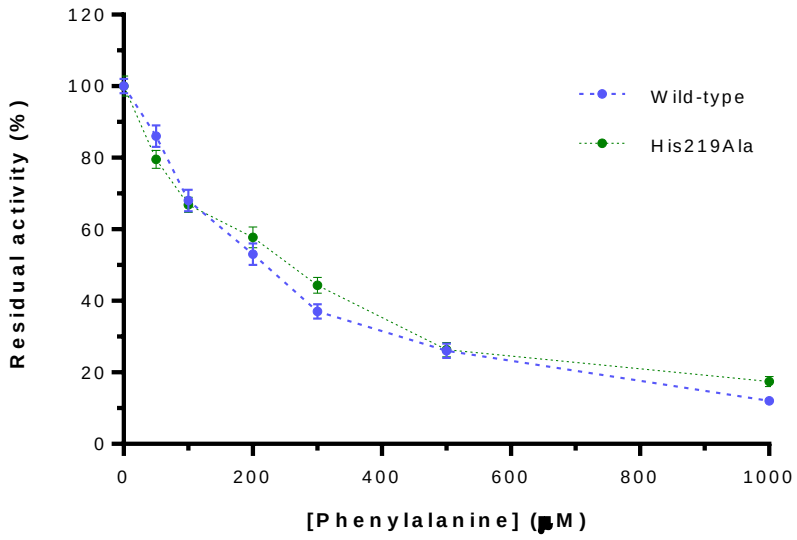


Figure 4.3.1.1 Inhibition studies of NmeDAH7PS^{His219Ala} in the presence of varying concentrations of allosteric effector, Phe. NmeDAH7PS^{His219Ala} (blue) and NmeDAH7PS^{WT} (green)

4.3.2 Isothermal titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were employed to assess if substitution of His219 has an effect on the binding of the allosteric inhibitor, Phe. It is evident that the binding isotherm can be fit to a two-site sequential binding model, which is the same as *Nme*DAH7PS^{WT}. However, the affinity of Phe for both sites has decreased. The K_D for binding at the first site of this variant enzyme is 4 ± 1 µM which indicates similar affinity to the *Nme*DAH7PS^{WT}. The second binding event is weaker than the first; following the same trend as *Nme*DAH7PS^{WT}, with a K_D of 32 ± 2 µM (Figure 4.3.2.1C).

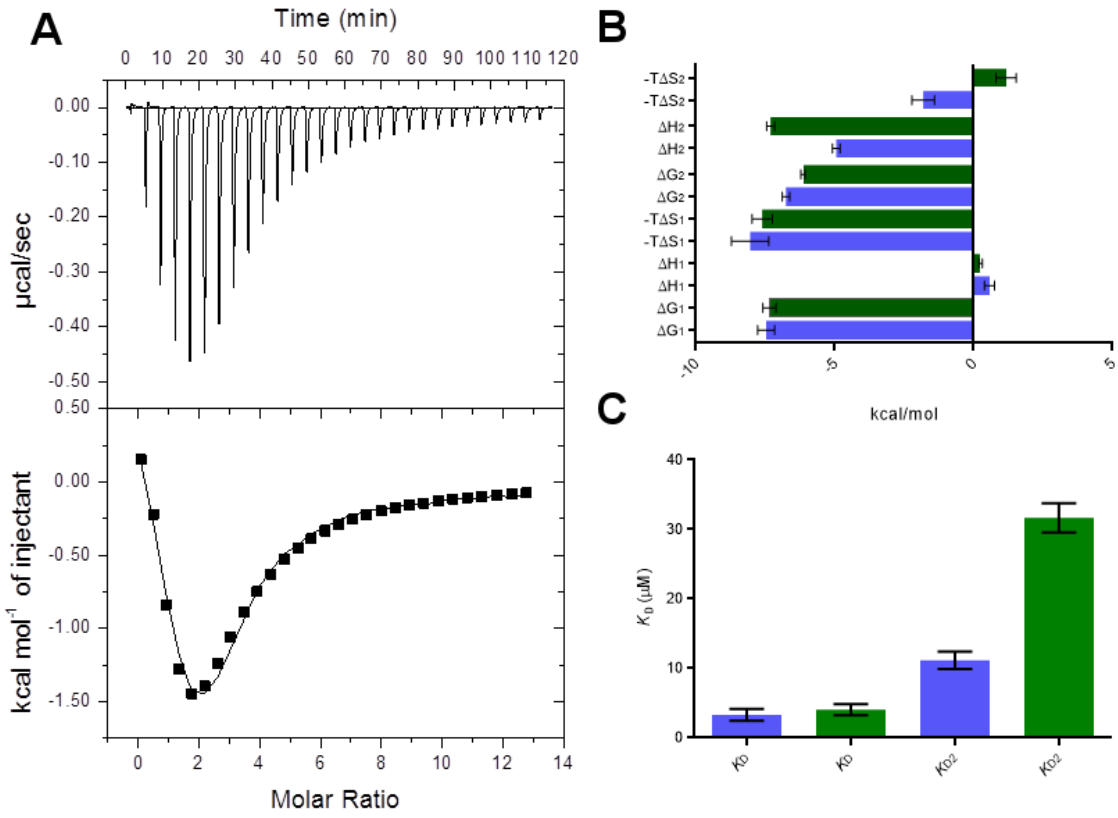


Figure 4.3.2.1 A) Raw data from ITC experiment involving the titration of 1.2 mM Phe into 20µM NmeDAH7PS^{His219Ala} variant, illustrating fit to the two-site sequential binding model. B) Comparison of the thermodynamic parameters for binding at both sites NmeDAH7PS^{WT} (blue) and NmeDAH7PS^{His219Ala} (dark green). C) Comparison of dissociation constant (K_D) at both sites of the sequential binding model, NmeDAH7PS^{WT} (blue) and NmeDAH7PS^{His219Ala} (dark green).

This binding isotherm is similar to the *Nme*DAH7PS^{WT} in that the binding at the first site decreases the affinity at the second site, which is in agreement with the negative cooperativity identified previously for this *Nme*DAH7PS^{WT}. However, inspection of the thermodynamic parameters indicates that binding of Phe at the second site is entropically

unfavourable in the *Nme*DAH7PS^{His219Ala} variant and this may contribute to the greater K_D for Phe binding at the second site compared to *Nme*DAH7PS (Figure 4.3.2.1B).

4.3.3 Analytical size exclusion chromatography

Analysis of the quaternary structural assembly of the enzyme *Nme*DAH7PS^{His219Ala} and *Nme*DAH7PS^{WT} was carried out by analytical size exclusion chromatography (SEC). The *Nme*DAH7PS^{His219Ala} variant eluted at a similar volume to *Nme*DAH7PS^{WT}, indicating, that it adopts the same tetrameric structure in solution (Figure 4.3.3.1).

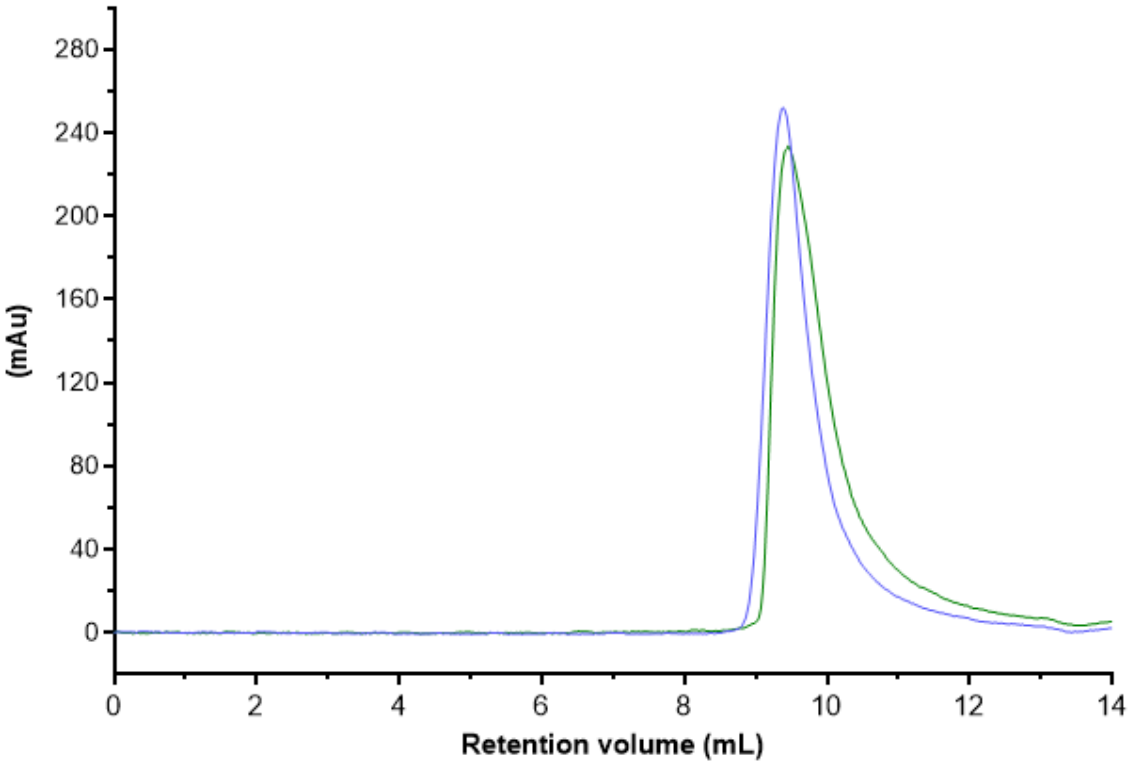


Figure 4.3.3.1 Analytical size exclusion chromatography, indicating the similarity on quaternary structure between NmeDAH7PS^{His219Ala} (dark green) and NmeDAH7PS^{WT} (blue) at 0.7 mg/mL.

4.3.4 Crystallography

Two crystal structures of the *Nme*DAH7PS^{His219Ala} variant were determined by X-ray diffraction. The apo structure was refined to a resolution of 2.21 Å and crystallised in the same space group as *Nme*DAH7PS^{WT} ($P12_11$) (PDB code: 4HSN). The Phe-bound structure crystallised in the same monoclinic space group and was refined to a resolution of 2.00 Å. The asymmetric

unit for both *Nme*DAH7PS^{His219Ala} structures contains a homotetramer made up of two tight dimers. Data collection and refinement statistics are shown in Table 4.3.4.1.

Crystal structures of the *Nme*DAH7PS^{His219Ala} variant in the presence and absence of Phe indicate that the enzymes structure is very similar to the *Nme*DAH7PS^{WT}. RMSD plots per residue indicate that the tight dimer remains largely unaffected by the His to Ala substitution and does not show any regions of significant variation (Figure 4.3.4.1).

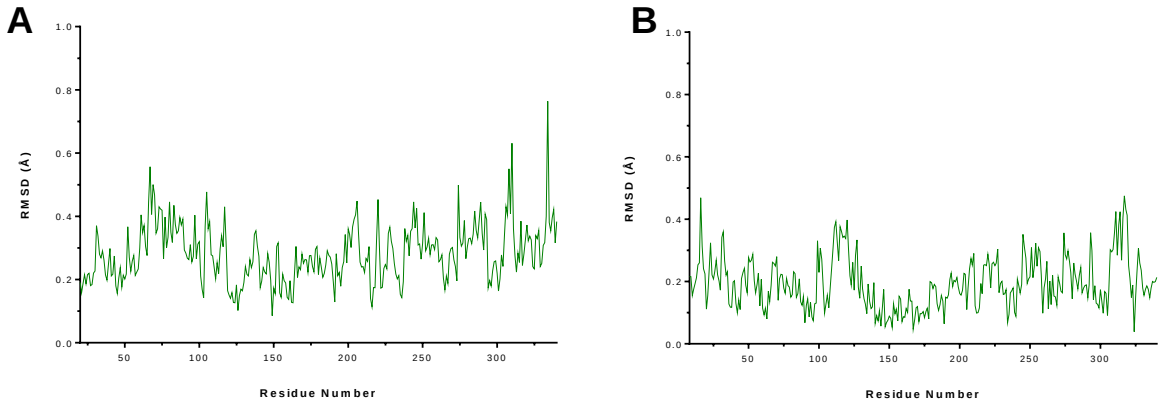


Figure 4.3.4.1 A) RMSD plot per residue between the PEP-bound NmeDAH7PS^{WT} structure (PDB code: 4HSN) and the apo NmeDAH7PS^{His219Ala}. B) RMSD plot per residue between the Phe-bound NmeDAH7PS^{WT} structure (PDB code: 4UC5) and the Phe-bound NmeDAH7PS^{His219Ala}

Alignment of the whole structures indicates that comparison of the tetrameric unit of the apo form of *Nme*DAH7PS^{His219Ala} to the PEP-bound *Nme*DAH7PS^{WT} (PDB code: 4HSN) yielded an RMSD for the alignment of 1252 Cα atoms of 0.349 Å. Comparison of the Phe-bound *Nme*DAH7PS^{His219Ala} to the Phe-bound *Nme*DAH7PS^{WT} (PDB code: 4UC5) yielded an RMSD for the alignment of 1368 Cα atoms of 0.260 Å. Inspection of the $|F_o\text{-}F_c|$ map for the Phe ligand as well as the temperature factors indicate that Phe is present in the allosteric site at full occupancy. This indicates that the substitution of these central histidines has minimal effect on the structure of *Nme*DAH7PS and its allosteric binding site (Figure 4.3.4.2). There is no variation in the dihedral angle between tight dimers in the *Nme*DAH7PS^{His219Ala} variant compared to the *Nme*DAH7PS^{WT}, indicating that the interface interactions are conserved in this variant. However, the similarity observed for this structure may arise from the network of water molecules located at the tetramer interface in both variant structures in the absence of these residues (Figure 4.3.4.2).

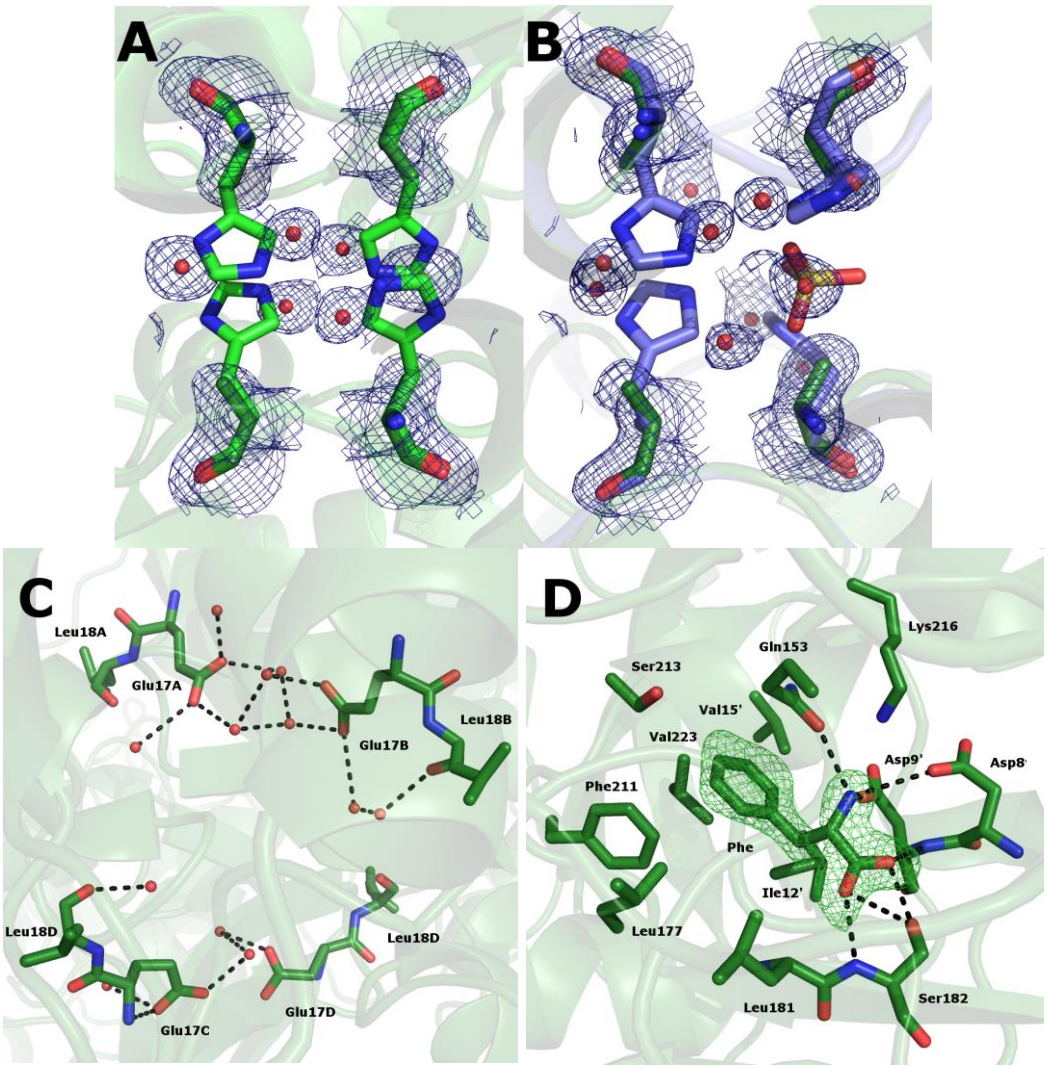


Figure 4.3.4.2 Occupation of the central region of the tetramer by a network of water molecules upon removal of Hi219 side-chain. A) $|2F_o\text{-}F_c|$ map contoured at 1 σ comparing the PEP-bound NmeDAH7PS^{His219Ala} variant to the PEP-bound NmeDAH7PS^{WT} structure (PDB code: 4HSN) B) $|2F_o\text{-}F_c|$ map contoured at 1 σ comparing the Phe-bound NmeDAH7PS^{His219Ala} variant to the Phe-bound NmeDAH7PS^{WT} structure (PDB code: 4UC5). C) Network of water molecules interacting with residues Glu17 and Leu18 of the N-terminal binding loop in the Phe-bound NmeDAH7PS^{His219Ala} structure. D) $|F_o\text{-}F_c|$ map contoured at 3 σ indicating presence of Phe at the allosteric site at full occupancy.

Superimposition of the tetramer for *Nme*DAH7PS^{His219Ala} and *Nme*DAH7PS^{WT} for both the Phe-bound and PEP-bound structures, indicates that the mutation does not have much of an effect on the substituted residue. The position of Cα residue 219 remains in the same position. This implies that interactions at the tetramer interface are not dependent on this residue in the crystalline form. This could however be due to the replacement of His219 by a network of water molecules upon mutation to alanine that do not exist in either PEP- or Phe-bound *Nme*DAH7PS^{WT} structures. In the presence of Phe, when His219 adopts its secondary conformation, it forms a hydrogen bond with Glu17 and the backbone carbonyl of Leu18. The

water molecules present instead of His219 in the Phe-bound *Nme*DAH7PS^{His219Ala} variant crystal structure may contribute to the closure of the allosteric binding loop when Phe binds at the allosteric site. This may be due to the ability of these water molecules to mimic certain characteristics of the His219 residue by forming hydrogen bonds with Glu17 and Leu18 of the N-terminal extension.

***Nme*DAH7PS^{His219Ala}**		
	L-Phenylalanine	**PEP**
Data Collection		
Crystal system; space group	Monoclinic, $P12_11$	Monoclinic, $P12_11$
Unit cell parameters		
a, b, c (Å)	73.45, 142.12, 74.92	73.72, 134.10, 76.13
α, β, γ (º)	90, 96.58, 90	90, 96.88, 90
Resolution range (Å)	74.43-2.00 (2.03-2.00)	75.58-2.21 (2.26-2.21)
Measurements	755228	579636
Unique reflections	101790	75734
Redundancy	7.4 (6.6)	7.7 (7.6)
Completeness (%)	99.1 (95.9)	100.0 (100)
I/σ (I)	15.2 (2.2)	19.1 (3.0)
R_{merge}	0.091 (0.813)	0.088 (0.700)
$CC_{1/2}$	0.710	0.779
Wilson B value (Å)^2	22.19	32.16
Matthews coefficient	2.55 (51.77)	2.54 (51.65)
Refinement		
R_{work}	0.2041	0.1910
R_{free}	0.2543	0.2437
Chain length	351	351
Observed number of residues	345 (Chain A, B & C) 347 (Chain A)	334 (Chain A), 332 (Chains B & D), 330 (Chain C)
Water molecules	730	384
Other	8 (Mn^{2+}, SO_4^{2-}, Cl^-)	8 (Mn^{2+}, SO_4^{2-}, EDO, PEG, Cl^-)
Ligand	4 (Phe), 1 (PEP)	2 (PEP)
Mean *B* (Å)^2		
Protein	32.26	46.68
Water	35.37	41.47
Other	38.03	59.88
Ligand	24.98, 36.69	42.24
R.m.s.d from target values		
Bond lengths (Å)	0.0123	0.0113
Bond angles	1.5065	1.4226
Dihedral angles	0.0852	0.0763
Ramachandran		
Preferred (%)	96.59	96.30
Allowed (%)	3.05	3.40
Outliers (%)	0.36	0.30
PDB Entry	-	-

Table 4.3.4.1 Data collection and refinement statistics of NmeDAH7PS^{His219Ala} variant in the presence and absence of Phe

4.4 DISCUSSION

The *Nme*DAH7PS^{His219Ala} variant shows no sign of significant variation from the *Nme*DAH7PS^{WT} enzyme in any of the experimental methods used for characterisation. It remains tetrameric in

solution, indicating that this single mutation is not sufficient to disrupt the quaternary structure. However, the specific activity of this variant has reduced significantly upon mutation. This is consistent with what was observed for the *Nme*DAH7PS^{Arg126Ser} variant (Part I) and implies that interactions at the tetramer interface affect catalytic turnover.

The inhibition of *Nme*DAH7PS^{His219Ala} by Phe is the same as *Nme*DAH7PS^{WT}, and the negative cooperativity observed between binding sites is consistent with the *Nme*DAH7PS^{WT}. The reduced affinity of the second allosteric site of the two-site sequential model obtained from the ITC binding isotherm for Phe, may be due to the inability to form hydrogen bonds between the sidechain of Glu17 and the backbone carbonyl of Leu18 of the N-terminal extension with residue 219 in the absence of the His219 sidechain. This interaction is observed in the Phe-bound *Nme*DAH7PS^{WT} structure (PDB code: 4UC5). These hydrogen bonds may aid in the closure of the N-terminal extension over the bound ligand. However, this feature does not have an effect on the inhibition caused by Phe at any concentration obtained from kinetic inhibition assays.

The crystal structures of the Phe-bound and apo forms of *Nme*DAH7PS^{His219Ala} indicate that there is a significant increase in the number of water molecules at the interface, which may replace the His side-chains in the *Nme*DAH7PS^{His219Ala} variant. The similarity between crystal structures of the *Nme*DAH7PS^{His219Ala} variants and *Nme*DAH7PS^{WT} is unsurprising as crystal packing effects govern the interactions seen at these notoriously weak interfaces, like that seen in the dimeric *Nme*DAH7PS^{Arg126Ser} variant (PDB code: 4UCG), which crystallised as a tetramer.

Overall, this residue is not observed to be crucial for the functionality of *Nme*DAH7PS. This provides further evidence for the lack of importance of the interactions at the interface for enzyme function. Taken together with what was observed for the *Nme*DAH7PS^{Arg126Ser} variant, the necessity to form this interface in *Nme*DAH7PS is negligible. This may have direct implications for the evolution of a tetrameric interface in other type Iα enzymes.

CHAPTER 5- EXPOSING THE COMMUNICATION OF ALLOSTERIC SIGNAL IN *NME*DAH7PS VIA THE CHARACTERISATION OF SEVERAL VARIANT ENZYMES

5.1 INTRODUCTION

Having collected a crystal structure of the Phe-bound form of *Nme*DAH7PS, described in chapter 3, attempts were made to identify structural changes upon Phe binding, with a view to attempt to elucidate a communication pathway between allosteric and active sites. While there is no observable major conformational change in *Nme*DAH7PS upon Phe binding in the crystalline form, it is evident that there is a subtle reorganisation of the hydrogen bonding network that connects allosteric and active sites in this enzyme. This reorganisation is believed to be involved in the reduction of catalytic efficiency and the affinity of the active site for both substrates.[26,123] Removal of the redundancy of the hydrogen bonds present in the PEP-bound structure (PDB code: 4HSN) and comparison with those from the Phe-bound structure (PDB code: 4UC5) reveals the hydrogen bonds that are broken and formed upon Phe binding to *Nme*DAH7PS in the crystalline form. Focusing only on the functional tight dimer of *Nme*DAH7PS, it is evident that the reorganisation is centred around the $\beta_2\alpha_2$ loop and its interaction with surrounding residues at the tight dimer interface, as well as the N-terminal allosteric binding loop as it becomes ordered upon Phe binding (Figure 5.1.1 A and B).

Chapter three described how the binding of Phe at the allosteric site caused a change in the conformation of the $\beta_2\alpha_2$ loop. This may be related to the reorganisation of the electrostatic network of residues on and surrounding the loop. The reorganisation of this network is intimately associated with the binding of Phe at both sites of the *Nme*DAH7PS tight dimer. Careful inspection of this reorganisation upon Phe binding, highlighted several ionisable residues that may be involved in the communication of allosteric signal, both within the same chain as well as with the adjacent chain across the tight dimer interface (Figure 5.1.1 C). In the PEP-bound crystal structure (PDB code: 4HSN), Glu98 is connected to the allosteric site of the same chain via hydrogen bonds with Tyr155, Thr151 and the backbone peptide amino group of Asp148, located on the $\beta_3\alpha_3$ loop and α_3 helix respectively. Lys107 forms interactions across dimer interface with the backbone carbonyl of His210′ and the side-chain of His209′ and Glu176′. However, in the Phe-bound structure (PDB code: 4UC5), the local environment of these residues as well as others on the $\beta_2\alpha_2$ loop changes dramatically. It is likely that this

reorganisation is due to the allosteric signal transmitted from the allosteric site upon Phe binding.

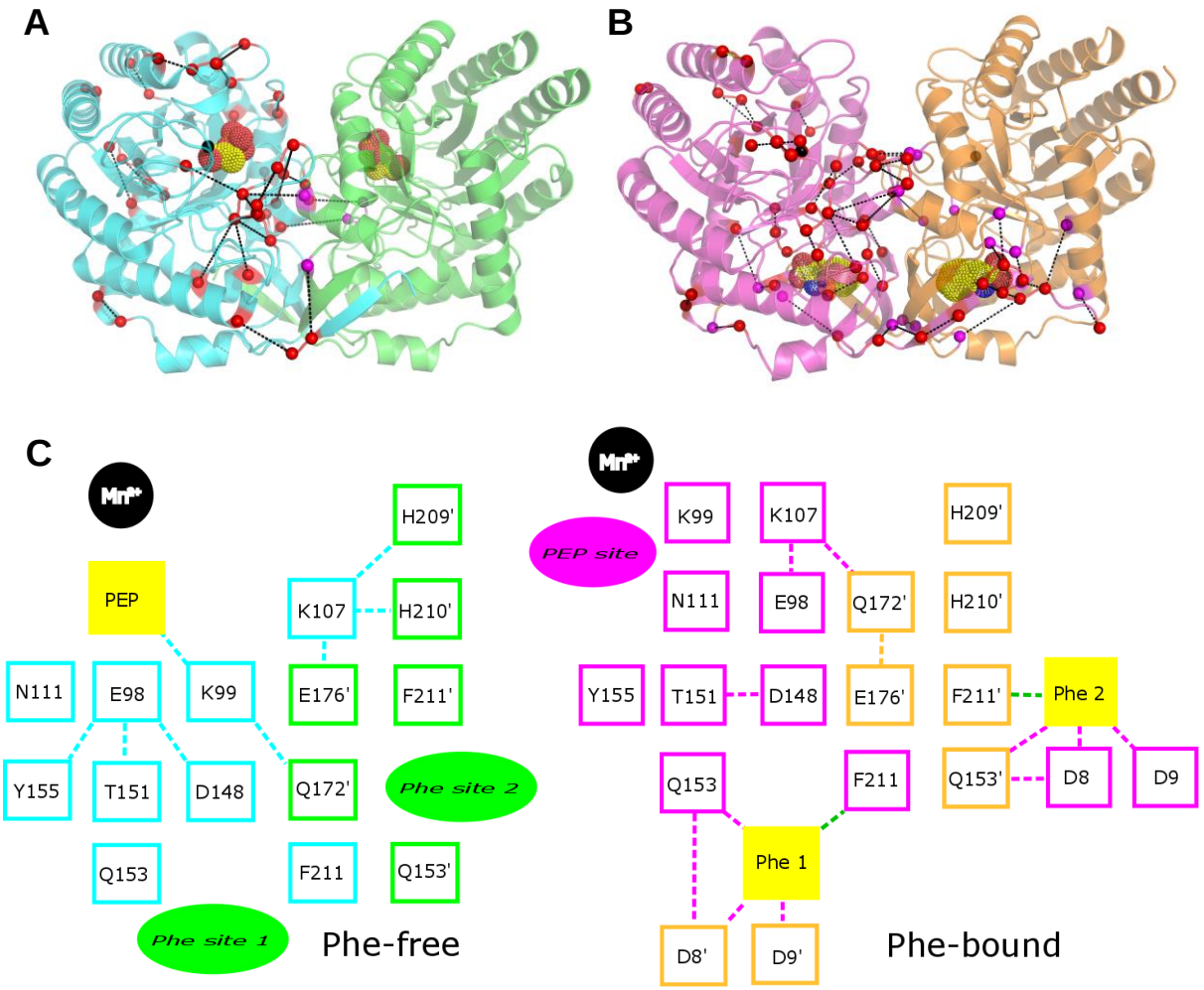


Figure 5.1.1 A) PEP-bound NmeDAH7PS (PDB code:4HSN) indicating hydrogen bonds broken upon Phe binding; Chain A (cyan), Chain B (green).B) Phe-bound NmeDAH7PS indicating hydrogen bonds formed upon Phe binding; Chain A (magenta), Chain B (Orange). Cα depicted as spheres, Chain A (red), Chain B (magenta). Mn^{2+} shown as black sphere. PEP and Phe shown as yellow dots. All hydrogen bonds are depicted as black dashed lines Hydrogen bonds obtained from PEP- and Phe-bound NmeDAH7PS structures (PDB code: 4HSN and 4UC5) using HBOND.[124] C) Schematic of hydrogen bond reorganisation upon Phe binding localised to the $\beta_2\alpha_2$ loop (coloured according to A and B).

Unfortunately, this analysis is not enough to confidently deduce a hydrogen bonding network linking remote sites. This is because we cannot discount the crystal packing effects, which may limit the conformations explored during the reorganisation of this network in the crystalline form, or the acquirement of an energy minimum by the protein that does not properly reflect what may occur in solution. This gave rise to a more in depth investigation into the transmission of this allosteric signal, with a view to properly define the role of each ionisable

residue in the transfer of this allosteric signal, through monitoring the pK_a variations over the course of a molecular dynamics (MD) simulation.

This work has been submitted for publication and is located in Appendix A6

This manuscript has been submitted to: Journal of the American Chemical Society

Heyes, L. C.*, Lang, E. J. M.*, Jameson. G. B. & Parker, E. J. pK_a variations expose allosteric communication pathways. *JACS* **XX,** XXX-XXX (XXXX).

(*) denotes co-first author

Prof G. B. Jameson provided useful insight into the manuscript. E. J. M. Lang performed the molecular dynamics simulations and pK_a analysis and contributed to the writing of the manuscript. Prof E.J. Parker designed experiments and contributed to the writing of the manuscript. The author of this thesis contributed to the writing of the manuscript and carried out the pH-dependent kinetic inhibition experiments, crystallography, crystal structure analysis and the pH dependent isothermal titration calorimetry experiments.

Allostery is the process by which enzyme activity is regulated by cellular metabolites has received considerable attention in this thesis. When allostery takes place without major conformational change, the mechanisms involved are often difficult to delineate. These long range interactions between remote sites of a protein involving minimal conformational change have been theorised to be mediated by changes in the frequencies and amplitudes of atomic motions (dynamics).[81] The allosteric mechanism of *Nme*DAH7PS involves changes in dynamics, making the mechanisms difficult to resolve. Previous studies involving the combination of biophysical methods, such as NMR, ITC, MD and single molecule FRET have been successful in assessing the role of structure and dynamics in allosteric cooperativity in a number of different systems.[88,125–130]

It has been hypothesised that interaction networks that are formed or broken contribute significantly to allosteric communication in proteins.[131] The subtle reorganisation of residues within these contact networks have both an enthalpic and entropic contribution.[132] The

sensitivity of these networks to a single amino-acid substitution has highlighted the importance of interactions between individual residues of these networks.[133] The work described in Appendix A6 details a novel approach to understanding these long range connections involving structural, MD and experimental studies. It involved monitoring the long range coulombic interactions, hydrogen bonding interactions and desolvation effects, which contribute to changes in pK_a value of ionisable residues in the protein (calculated using PROPKA).[134] A significant variation in calculated pK_a value implies a substantial change in the local electrostatic environment of a residue, associated with a reorganisation of the hydrogen bonding network within the protein, which can occur upon binding of an allosteric effector.[135]

The activity of *Nme*DAH7PS is allosterically controlled through subtle changes in dynamics and does not undergo a large conformational change, making the mechanism of signal transduction intriguing. The $\beta_2\alpha_2$ loop, located within mutual proximity of both active and allosteric sites of each tight dimer, has been shown to be one of the most flexible regions of the protein during MD simulations for both the Phe-bound and Phe-free state, but was found to be more flexible in the Phe-bound state. The loop is intrinsically associated with residues of the β_{6a} and β_{6b} sheets, which make up a large portion of the predominantly hydrophobic allosteric site at tight-dimer interface. ITC experiments have shown that binding of Phe to the allosteric site of *Nme*DAH7PS has a large entropic character that may contribute to the negative cooperativity between the allosteric sites of the enzyme.

Deep within the allosteric site, hydrophobic side-chains are abundant and have been shown to play a crucial role in recognition and selectivity of the allosteric aromatic ligand (Chapter 3). Interactions between the allosteric site hydrophobic residues and the allosteric ligand alter the interactions between the β-hairpin insertion and the $\beta_2\alpha_2$ loop across the tight dimer interface.

Several residues identified from structural and MD simulations have been proposed to play a role in communication of the allosteric signal between the active and allosteric site, all of which are intrinsically associated with the dynamics of the $\beta_2\alpha_2$ loop. Some of these interactions occur within the same chain, while others occur between chains, across the tight dimer interface. There are several residues that in the presence of Phe have been predicted

from the MD simulations to adopt a significantly altered pK_a value, in the presence or absence of an allosteric ligand. These residues are postulated to make up a hydrogen bonding network in the active enzyme and undergo significant reorganisation in the presence of Phe. This hydrogen bonding network is reorganised so as to reduce the rate of catalysis, and the apparent K_M's for both substrates.[26,123] Such a highly specific mechanism may have implications for the flux and control of the entire shikimate biosynthetic pathway in *N. meningitidis* and enable us to shed light on the allosteric mechanisms used in other organisms and enzymes from other biosynthetic pathways in nature.[20]

5.1.1 Hydrophobically driven binding event sets allostery in motion

The allosteric binding site of *Nme*DAH7PS largely consists of hydrophobic residues, which are involved in the recognition and binding of aromatic amino acids, in particular, Phe. The residues that make up this hydrophobic pocket are located at the interface of the two chains of the tight dimer. Phe211, Val214 and Val223 are from the β hairpin insertion, Leu177 and Leu181, located on the α_4 helix and $\alpha_4\beta_5$ loop and Ile12 and Val15 located on the N-terminal extension of the adjacent chain. Some of these residues have already been implicated in the allostery of the Phe sensitive isozyme from *E. coli.*[101] It has been shown that in the absence of the first 15 residues that make up the allosteric binding loop, allostery was abolished.[101] Phe 209 (211 in *Nme*DAH7PS), Val 221 (223 in *Nme*DAH7PS), Leu175 (177 in *Nme*DAH7PS) and Leu179 (181 in *Nme*DAH7PS) were also shown to be very important for allostery in *Eco*DAH7PS;[101] however, the reason for the loss of activity was not identified. Unfortunately, these experiments were largely inconclusive as there was no biophysical characterisation reported, nor any structural information to explain the loss of allostery seen in *Eco*DAH7PS.

Upon binding of Phe at the allosteric site, the N-terminal extension closes over the bound ligand and forms several hydrogen bonds via Asp8, Asp9, Gln153 and the backbone nitrogen of Ser182. Stabilisation of this closed conformation occurs when hydrogen bonds are formed between Asp8 of the allosteric binding loop, Gln153 on the α_3 helix and Lys216, which forms part of the $\beta_{6a}\beta_{6b}$ loop.

This study involves the generation of two point mutants (*Nme*DAH7PS^{Phe211Ala} and *Nme*DAH7PS^{Val223Ala}) and describes the structural, biophysical and functional characterisation of the influence of these residues in the allosteric mechanism of *Nme*DAH7PS.

5.1.2 Hydrogen bonding network reorganisation mediates allosteric signal transduction

The active site of *Nme*DAH7PS is predominantly made up of positively charged residues, which are involved in binding the negatively charged substrates. PEP binds deep within the active site, forming interactions with residues: Lys188, Arg236, Arg94 and Lys99. E4P is predicted to bind to the KPR(T/S) motif, part of the highly flexible $\beta_2\alpha_2$ loop, which has a role in mediating substrate binding and catalysis. The positioning of this loop is controlled by a network of residues, which have been shown to adopt different bonding partners in the presence and absence of Phe.

The residues identified as having potential roles in this network predicted from the structural and pK_a analysis described in chapter 5 are: Glu98, Lys99, Lys107, Asn111, Asp148, Thr151, Tyr155, Gln172′ and Glu176′, as well as a crystallographically conserved water molecule. The Cα of Glu98, located towards the N-terminal end of the $\beta_2\alpha_2$ loop moves ~2.5 Å upon binding of Phe. In the PEP-bound structure (PDB code: 4HSN), Glu98 forms hydrogen bonds with the backbone nitrogen of Asp148 and the side-chain hydroxyl groups of Thr151 and Tyr155, from the $\beta_3\alpha_3$ loop and α_3 helix. This loop contains Gln153, which is responsible for binding the amine moiety of Phe (Figure 5.1.2.1). In the Phe-bound *Nme*DAH7PS^{WT} structure (PDB code: 4UC5), Glu98 forms new hydrogen bonds with Gln172′ and Asn111, which was involved in a hydrogen bond with the backbone carbonyl of His210′ from the β_{6a} sheet in the PEP-bound enzyme (Figure 5.1.2.1). Glu98 also comes into proximity to form a salt-bridge with Lys107 and interacts with Glu176′ via a shared water molecule. This water molecule is not however conserved for the duration of MD simulations of the Phe-bound state of *Nme*DAH7PS^{WT}. When this reorganisation occurs upon Phe binding, it is not surprising that the positioning of the adjacent residue Lys99 is also affected. As this residue is part of the absolutely conserved KPR(T/S) motif, and has involvement in proton transfer during catalysis, its new conformation may contribute to the flipped binding mode of PEP, where the divalent active site metal ion is

no longer in close proximity to the carboxylate of PEP, nor accessible for coordination by the aldehyde of E4P.[22,28] The residue Lys107 is located towards the C-terminal end of the $\beta_2\alpha_2$ loop and in the PEP-bound *Nme*DAH7PS^{WT} structure. This residue forms a salt bridge across the tight dimer interface with Glu176′ and hydrogen bonds with His 209′ and the backbone carbonyl of His210′. In the presence of Phe, these interactions are lost, in part due to the movement of the $\beta_2\alpha_2$ loop and β_{6a} sheet, which contains Phe211′, believed to contribute to allosteric ligand recognition. In the presence of Phe, Lys107 forms a new hydrogen bond with Gln172′ and has been shown from MD simulations to interact with Glu98 directly, even if this interaction is > 3.4 Å in the crystal (Figure 5.1.2.1).

The aforementioned ionisable and polar residues are hypothesised to have a critical role in both catalysis and regulation. Systematic mutation of each residue presumably involved in this reorganisation upon binding of the allosteric effector molecule was completed in order to determine their involvement in the network. In doing so, improving our understanding of how the allosteric signal is propagated from the allosteric site in *Nme*DAH7PS, resulting in a functional response at the active site.

All variant enzymes were designed to remove the side-chain functionality of respective residues in order to determine their involvement in the mediation of allosteric signal. The mutated residues include: Glu98, Lys107, Gln172 and Glu176. Conversely, a variant was made based on the pK_a predictions, where Glu176′ was calculated to undergo a significant change in pK_a in the presence of Phe. In light of this observation, a variant was generated to enable the formation of hydrogen bonds within the hydrogen bonding network, but with the inability to undergo a change in pK_a. Generation of the *Nme*DAH7PS^{Glu176Gln} variant appropriately satisfied these criteria, as it was considered that this change may have a unique effect on the communication of allosteric signal.

The kinetic, thermodynamic and structural characterisation of these variant enzymes may help to elucidate the mechanism employed by *Nme*DAH7PS to achieve allostery and provide a general understanding of allostery in other type Iα DAH7PS from different organisms. These studies reveal how *Nme*DAH7PS is regulated without significant conformational change, extensive accessory domains, or protein complex formation, as is seen in various type Iβ and type II DAH7PS enzymes.[20]

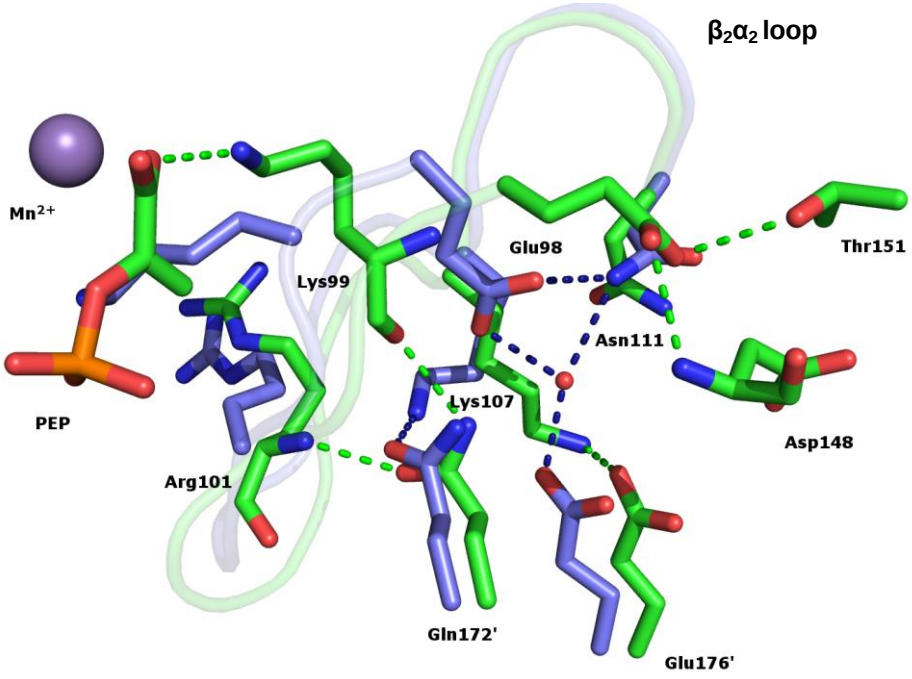


Figure 5.1.2.1. Hydrogen bonding network, in the presence of PEP (green) and in the presence of Phe (blue), obtained from structural alignment of the tight dimers of NmeDAH7PS (PDB codes: 4HSN and 4UC5). Hydrogen bonds indicated in green and blue respectively. The conserved water molecule from Phe-bound structure is shown as a red sphere.

5.2 RESULTS FROM ALLOSTERIC BINDING SITE VARIANTS

5.2.1. Kinetic characterisation of allosteric binding site variants

The kinetic parameters of the allosteric binding site *Nme*DAH7PS variant enzymes were determined under the conditions used for *Nme*DAH7PS^{WT} (section 7.1.27). The kinetic parameters obtained for these allosteric site variant enzymes are shown in Table 5.2.1.1. The *Nme*DAH7PS^{Phe211Ala} variant has a K_M for PEP of 6 ± 1 µM, which is lower than that of *Nme*DAH7PS^{WT}, indicating higher affinity of the active site for this variant for PEP. It also exhibits a ~2-fold increase in the K_M for E4P of 71 ± 6 µM. The k_{cat} for this variant was also reduced to ~60% of that observed for *Nme*DAH7PS^{WT}. The *Nme*DAH7PS^{Val223Ala} variant has similar kinetic properties to the *Nme*DAH7PS^{WT}, with a K_M for the substrates E4P and PEP of 36 ± 4 µM and 16 ± 2 µM respectively. The k_{cat} for this variant is also very similar to the *Nme*DAH7PS^{WT} with a value of 32 ± 0.2 s^{-1}.

	PEP K_M (µM)	E4P K_M (µM)	k_{cat} (s^{-1})	$k_{cat}/K_{M\ (PEP)}$	$k_{cat}/K_{M\ (E4P)}$
***Nme*DAH7PS^{Phe211Ala}**	6 ± 1	71 ± 6	15 ± 0.2	2.5	0.2
***Nme*DAH7PS^{Val223Ala}**	16 ± 2	36 ± 4	32 ± 0.2	2.0	0.9
***Nme*DAH7PS^{WT}**	15 ± 1	37 ± 2	27.1 ± 0.1	1.8	0.7

Table 5.2.1.1 Kinetic parameters of NmeDAH7PS^{Phe211Ala} and NmeDAH7PS^{Val223Ala} variants. Equivalent parameters for NmeDAH7PS^{WT} are given as a comparison. All kinetic assays were carried out at high concentrations of other substrate.

5.2.2. Allosteric binding site variants are less sensitive to Phe

Inhibition kinetics were carried out using standard assay procedures (section 7.1.27) in the presence of varying concentrations of Phe (50–1000 µM). Sensitivity to Phe was severely reduced for both allosteric site variant enzymes. *Nme*DAH7PS^{Phe211Ala} was found to be insensitive to inhibition by Phe at all concentrations tested. *Nme*DAH7PS^{Val223Ala} was inhibited by Phe, but the level of inhibition was decreased severely compared to *Nme*DAH7PS^{WT}, where 62% residual activity was maintained at 1 mM Phe, compared to the 12% retained for *Nme*DAH7PS^{WT} (Figure 5.2.2.1).

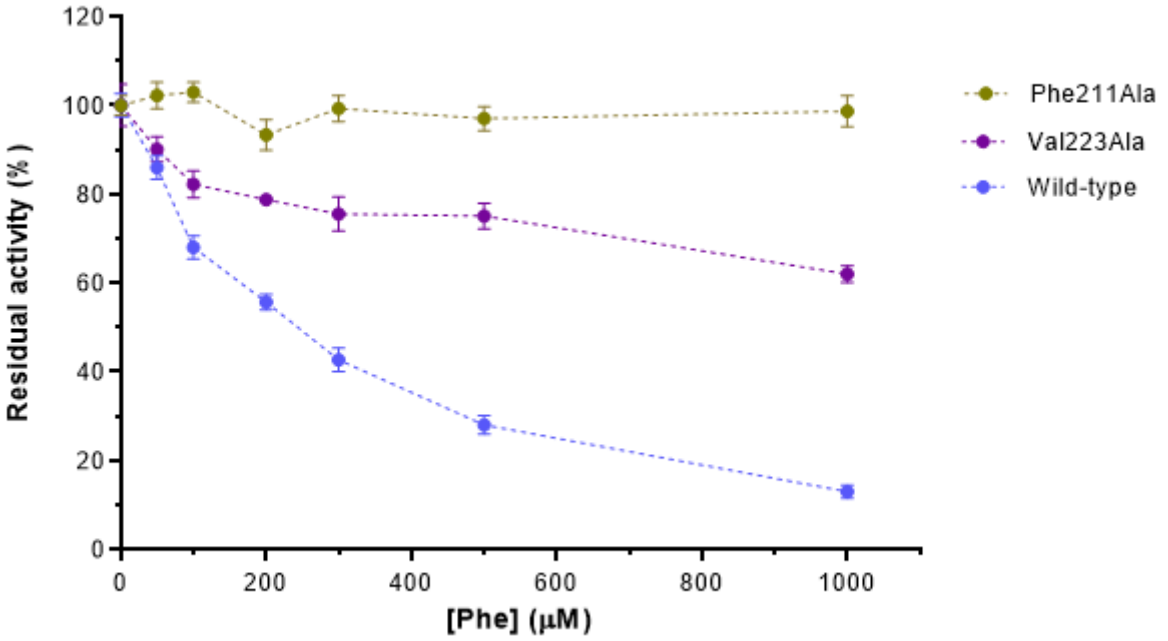


Figure 5.2.2.1 Inhibition of NmeDAH7PS^{Phe211Ala} variant (khaki) and NmeDAH7PS^{Val223Ala} variant (purple) compared to the NmeDAH7PS^{WT} enzyme (blue) at varying concentrations of phenylalanine (0-1000 µM). Saturating substrate concentrations were used for all experiments.

Standard ITC protocols were employed to elucidate the mechanism of decreased sensitivity to Phe observed for these variant enzymes (section 7.1.29). The insensitivity of *Nme*DAH7PS^{Phe211Ala} variant to high concentrations of Phe can be identified from the isotherm, where the protein did not undergo any heat change upon injection of Phe into the cell. Therefore, *Nme*DAH7PS^{Phe211Ala} no longer possesses the ability to bind Phe, implicating Phe211 directly in the binding of Phe at the allosteric site. The binding isotherm for the titration of Phe into *Nme*DAH7PS^{Val223Ala} is indicative of a very low affinity for Phe. The extremely high molar ratio obtained from this experiment renders the fit and calculated K_D meaningless. The shape of the curve may explain the reduction in inhibition seen for this variant enzyme compared to *Nme*DAH7PS^{WT}, which is a likely result of weaker binding at the allosteric site (Figure 5.2.2.2).

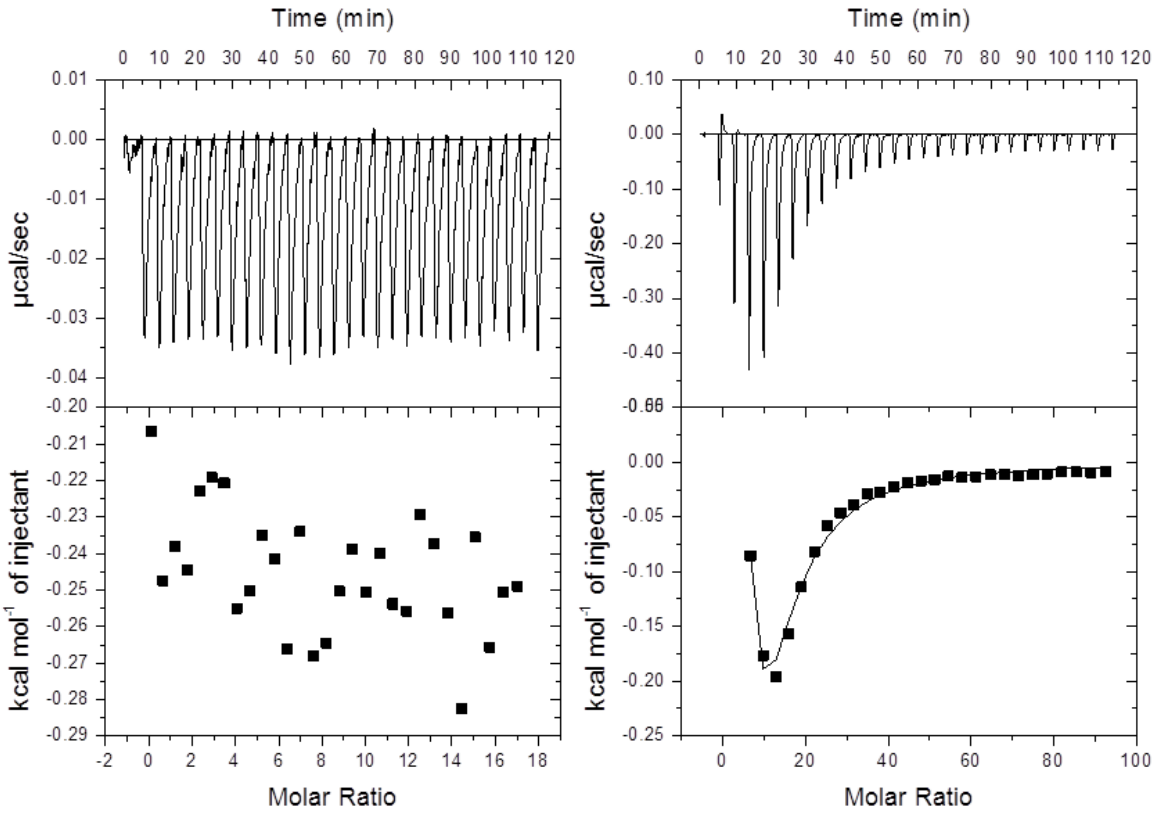


Figure 5.2.2.2 Raw data of the ITC experiments for the binding of Phe to NmeDAH7PS^{Phe211Ala} and NmeDAH7PS^{Val223Ala} Using the VP-ITC Microcal™ Left. NmeDAH7PS^{Phe211Ala} (1.2mM Phe and 16µM enzyme) Right. NmeDAH7PS^{Val223Ala} (20mM Phe and 20µM enzyme).

5.2.3. Allosteric binding site variants adopt tetrameric quaternary structure in solution

Analytical size-exclusion chromatography (SEC) was employed to identify whether the quaternary structure of these variant enzymes was altered upon introduction of respective mutations. The main peak for *Nme*DAH7PS^{Phe211Ala} and *Nme*DAH7PS^{Val223Ala} eluted from the column after the same volume of isocratic flow as *Nme*DAH7PS^{WT}. This indicates that they adopt the same tetrameric solution-state quaternary structure (Figure 5.2.3.1).

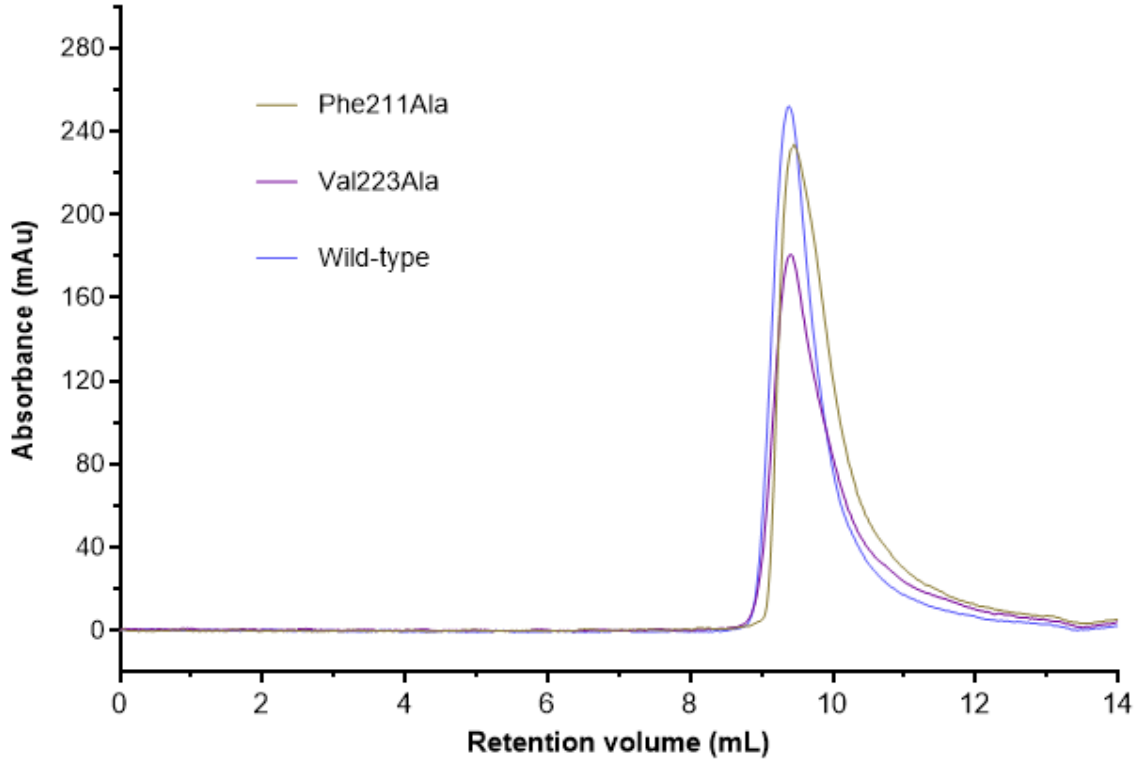


Figure 5.2.3.1 Analytical size exclusion chromatography traces of NmeDAH7PS^{WT} (blue), NmeDAH7PS^{Phe211Ala} variant (khaki), NmeDAH7PS^{Val223Ala} variant (purple). Concentrations used for the experiment varied from 0.6-0.7 mg/mL.

5.2.4. Crystallography of allosteric binding site variants

In order to identify structural differences of these two allosteric binding site variant enzymes that may contribute to their lessened affinity for Phe, crystal structures of both *Nme*DAH7PS^{Phe211Ala} and *Nme*DAH7PS^{Val223Ala} were obtained by X-ray diffraction methods. Crystals of these variant proteins were grown using procedures described in section 7.1.30

and 7.1.31. The soaking of Phe into these structures was achieved using procedures described in 7.1.32. Data collection and refinement statistic are shown in table 5.2.4.1.

***Nme*DAH7PS**	***Val223Ala***	***Val223Ala***	***Phe211Ala***
	L-Phenylalanine	**PEP**	**-**
Data Collection			
Crystal system; space group	Orthorhombic, $P2_12_12_1$	Monoclinic, $P12_11$	Orthorhombic, $P2_12_12_1$
Unit cell parameters			
a, b, c (Å)	73.62, 75.43, 143.94	73.26, 135.62, 75.99	78.68, 135.73, 149.36
α, β, γ (º)	90, 90, 90	90, 96.27, 90	90, 90, 90
Resolution range (Å)	75.43-1.70 (1.74-1.70)	75.53-1.84 (1.88-1.84)	100.45-2.35 (2.41-2.35)
Measurements	645375	480663	964977
Unique reflections	88844	126857	67384
Redundancy	7.3 (7.3)	3.8 (3.8)	14.3 (14.8)
Completeness (%)	100.0 (100.0)	99.6 (99.3)	100.0 (100)
I/σ (I)	14.5 (2.27)	13.4 (1.66)	19.4 (2.51)
R_{merge}	0.095 (0.769)	0.054 (0.942)	0.104 (0.986)
$CC_{1/2}$	0.763	0.531	0.909
Wilson B value $(Å)^2$	16.10	25.26	42.10
Matthews coefficient	2.63 (53.25)	2.47 (50.2)	2.63 (53.24)
Refinement			
R_{work}	0.1516	0.1623	0.2097
R_{free}	0.1782	0.1970	0.2589
Chain length	351	351	351
Observed number of residues	345 (Chain A & B)	332 (Chain A, B, C & D)	333 (Chain A+B), 332 (Chain C), 328 (Chain D)
Water molecules	613	733	333
Other	18 (Mn^{2+}, Cl^-, EDO, PEG)	12 (Mn^{2+}, SO_4^{2-}, Cl^-, EDO, PEG)	9 (PEG, EDO, Mn^{2+})
Ligand	2 (Phe)	4 (PEP)	1 (PEP)
Mean *B* $(Å)^2$			
Protein	17.66	32.80	44.76
Water	26.82	36.18	42.15
Other	47.78	59.43	65.14
Ligand	14.51	30.83	63.06
R.m.s.d from target values			
Bond lengths (Å)	0.0121	0.0130	0.0116
Bond angles	1.4671	1.5248	1.4850
Dihedral angles	0.0888	0.0912	0.0814
Ramachandran			
Preferred (%)	97.11	97.66	96.38
Allowed (%)	2.65	2.19	3.39
Outliers (%)	0.14	0.15	0.23
PDB Entry	5D04	5D03	5D09

Table 5.2.4.1 Data collection and refinement statistics of NmeDAH7PS^{Val223Ala} variant in the presence and absence of Phe and the NmeDAH7PS^{Phe211Ala} variant (PDB code: 5D03, 5D04 & 5D09)

Unfortunately, a Phe-bound structure of the *Nme*DAH7PS^{Phe211Ala} variant enzyme could not be obtained, which correlates to the lack of evidence for Phe binding or inhibition from ITC and kinetic experiments. However, at the high concentrations (>5 mM) used for soaking Phe into crystals of *Nme*DAH7PS^{Val223Ala} variant, a structure was obtained with Phe-bound at the allosteric site at full occupancy.

Crystal structure of apo NmeDAH7PS^{Phe211Ala}

A single crystal structure of the apo *Nme*DAH7PS^{Phe211Ala} variant was determined by X-ray diffraction. The structure was refined to a resolution of 2.35 Å and crystallised in an orthorhombic space group, ($P2_12_12_1$). Inspection of the structure highlights some differences to the *Nme*DAH7PS^{WT} structure. Alignment of the tight dimers of the *Nme*DAH7PS^{Phe211Ala} variant to the PEP-bound *Nme*DAH7PS^{WT} (PDB code: 4HSN) yielded an RMSD of 0.387-0.494 Å. The regions of most significant variation were localised to the highly flexible $\beta_2\alpha_2$ loop, which adopts a different conformation to what was observed for the PEP-bound *Nme*DAH7PS^{WT} (Figure 5.2.4.1A). The reason for the absence of PEP from this structure remains unclear, but may be an artefact of the crystallisation procedure. It is important to bear this in mind when considering the structural changes associated with this mutation.

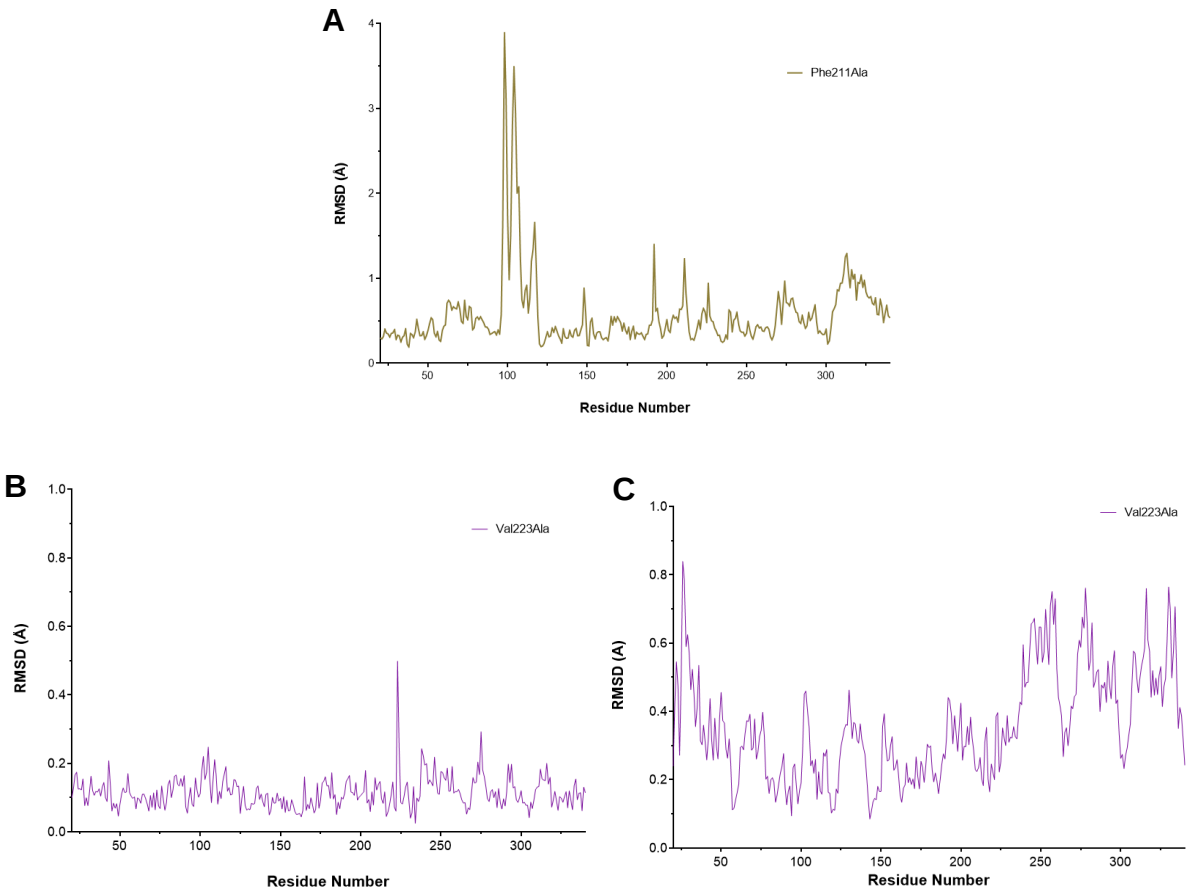


Figure 5.2.4.1 Root mean squared deviation of the tight dimer averaged for both chains. A) Comparison between PEP-bound NmeDAH7PS^{Phe211Ala} (khaki) and PEP-bound NmeDAH7PS^{WT} (PDB code: 4HSN). B) Comparison between PEP-bound NmeDAH7PS^{Val223Ala} (purple) and PEP-bound NmeDAH7PS^{WT} (PDB code: 4HSN). C) Comparison between Phe-bound NmeDAH7PS^{Val223Ala} (purple) and Phe-bound NmeDAH7PS^{WT} (PDB code: 4UC5).

The $\beta_2\alpha_2$ loop (res. 95-120) responsible for binding E4P illustrates an interesting difference between the apo-*Nme*DAH7PS^{Phe211Ala} and PEP-bound *Nme*DAH7PS^{WT}. Asn111 in the PEP-bound *Nme*DAH7PS^{WT} structure forms a hydrogen bond with the backbone carbonyl of His210' from the adjacent chain of the tight dimer. However, upon mutation of Phe211 to alanine, the distance between Asn111 and His210' is too large for the formation of a hydrogen bond. The loss of this structural stabilisation may contribute to the RMSD between the $\beta_2\alpha_2$ loops of these structures. The variation in conformation of this loop, may also clarify the decreased binding affinity for E4P obtained from kinetic studies.

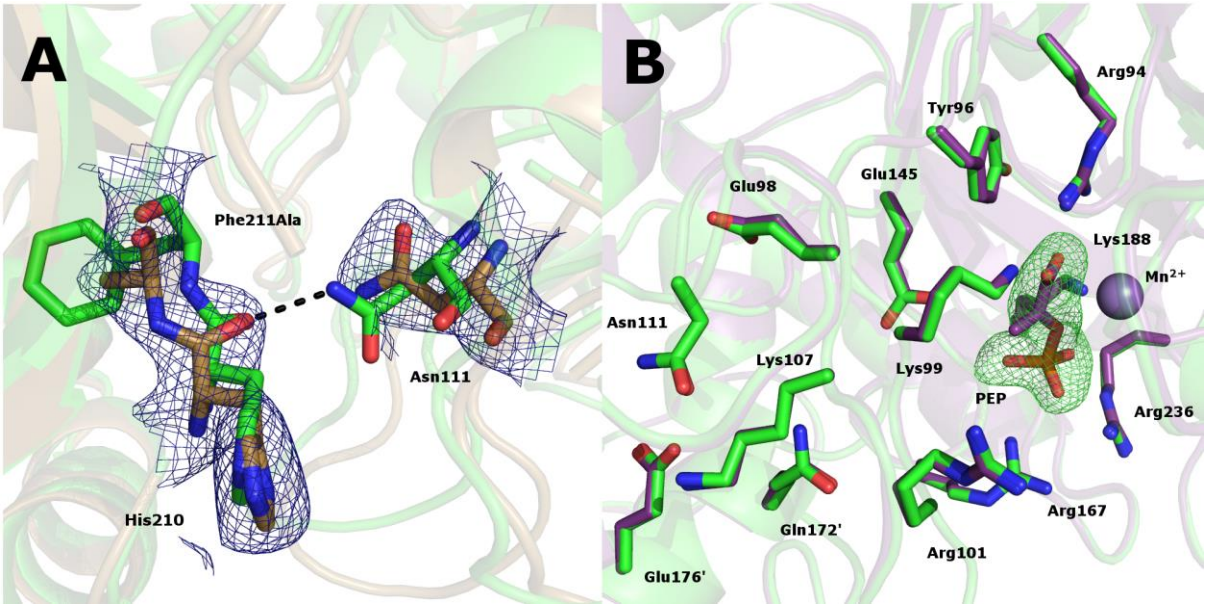


Figure 5.2.4.2. A) Structural alignment of NmeDAH7PS^{Phe211Ala} variant and PEP-bound NmeDAH7PS^{WT} (PDB code: 4HSN) illustrating the loss of hydrogen bond at the tight dimer interface. $|2F_o\text{-}F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate mutation. B) Structural alignment of the PEP-bound NmeDAH7PS^{Val223Ala} variant and PEP-bound NmeDAH7PS^{WT} (PDB code: 4HSN) illustrating the conservation of residue positions at the active site. PEP $|F_o\text{-}F_c|$ omit map coloured in green and contoured at 3.0 σ.

Crystal structures of Phe- and PEP-bound NmeDAH7PS^{Val223Ala}

Two crystal structures of the *Nme*DAH7PS^{Val223Ala} variant were determined by X-ray diffraction methods, one PEP-bound and the other Phe-bound. The PEP-bound structure was refined to a resolution of 1.84 Å and crystallised in the same space group as the PEP-bound *Nme*DAH7PS^{WT} (PDB code: 4HSN) ($P12_11$). The RMSD for the alignment of the tight dimers of the PEP-bound *Nme*DAH7PS^{Val223Ala} variant to the PEP-bound *Nme*DAH7PS^{WT} structure was 0.119-0.237 Å. These structures are remarkably similar (Figure 5.2.4.1B) with no observed regions of structural variation. This is reiterated for the PEP binding site and $\beta_2\alpha_2$ loop, which remain in

the same conformation when comparing the two enzymes (Figure 5.2.4.3B). It is perhaps not surprising that the kinetic parameters obtained for the *Nme*DAH7PS^{Val223Ala} variant are similar to the *Nme*DAH7PS^{WT}.

The Phe-bound structure crystallised in an orthorhombic space group ($P2_122_1$) and was refined to a resolution of 1.70 Å. The asymmetric unit for the Phe-bound *Nme*DAH7PS^{Val223Ala} structure (PDB code: 5D04) only contained the tight dimer as opposed to the tetramer, seen in all other structures, which may be due to the extensive twist of the tight dimers relative to each other for this variant compared to the Phe-bound *Nme*DAH7PS^{WT} (with an inter-dimer angle increase of ~10°). The RMSD for the alignment of the tight dimers of the Phe-bound *Nme*DAH7PS^{Val223Ala} variant to the Phe-bound *Nme*DAH7PS^{WT} structure (PDB code: 4UC5) was 0.308-0.339 Å, indicative of minimal variation between the tight dimers in these structures.

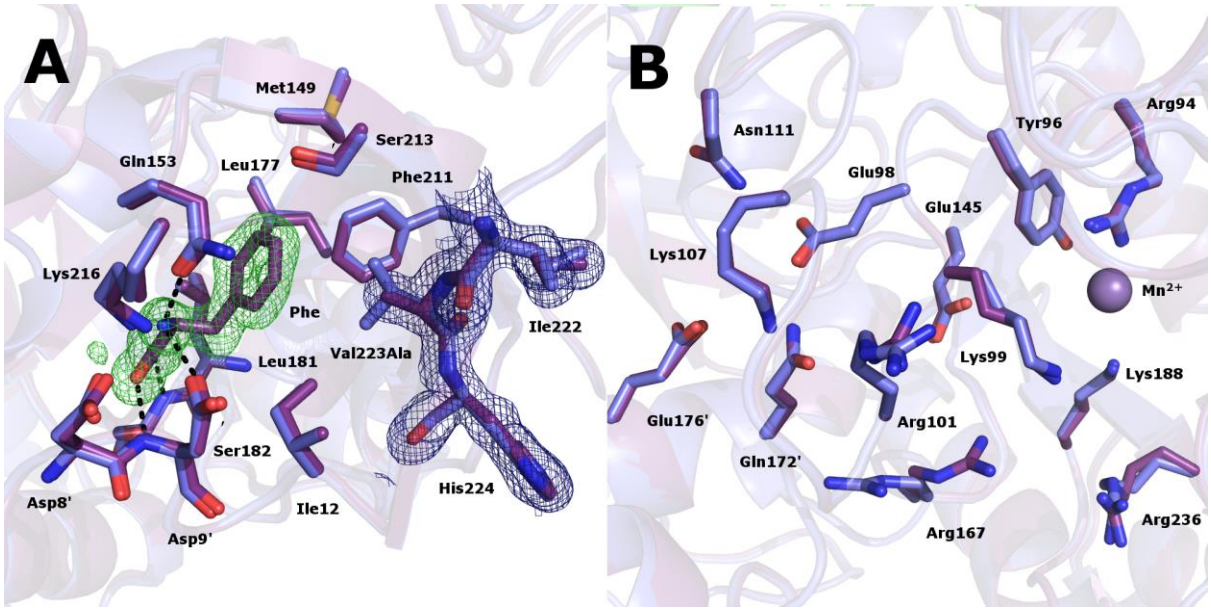


Figure 5.2.4.3. A) Structural alignment of the Phe-bound NmeDAH7PS^{Val223Ala} variant and Phe-bound NmeDAH7PS^{WT} (PDB code: 4UC5) illustrating the conservation of residue positions at the allosteric site. Phe $|F_o-F_c|$ omit map coloured in green and contoured at 3.0 σ and $|2F_o-F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate mutation. B) Structural alignment of the Phe-bound NmeDAH7PS^{Val223Ala} variant and Phe-bound NmeDAH7PS^{WT} (PDB code: 4HSN) active site.

Interestingly, the central histidines (His219) (described in chapter 4) in the Phe-bound *Nme*DAH7PS^{Val223Ala} structure engage in the same π-stacking interaction across the diagonal interface, and the Arg126-Glu27 salt bridge remains present in this structure. This indicates that the important interface contacts remain, even when the tight dimers are rotated by ~10°.

Inspection of the mutation site indicates no change in the position of the β_{6b} sheet on which this mutation is located and more specifically, there is no change in the position of Cα for Ala223 (Figure 5.2.4.3A). This implies that other residues within the site can facilitate Phe binding in the same way as is seen in *Nme*DAH7PS^{WT} (PDB code: 4UC5) at high concentrations used for Phe soaking.

5.3 RESULTS FROM HYDROGEN BONDING NETWORK VARIANTS

5.3.1 Kinetics and inhibition of hydrogen bonding network variants

Kinetic assays were carried out according to procedures described in (section 7.1.27). Characterisation of these hydrogen bonding network variant enzymes indicated some interesting differences in kinetic parameters when compared to *Nme*DAH7PS^{WT}. All kinetic constants are shown in table 5.3.1.1.

	PEP K_M (μM)	E4P K_M (μM)	k_{cat} (s^{-1})	$k_{cat}/K_{M\ (PEP)}$	$k_{cat}/K_{M\ (E4P)}$
***Nme*DAH7PS^{Glu98Ala}**	2 ± 0.2	29 ± 2	12 ± 0.2	6.0	0.4
***Nme*DAH7PS^{Lys107Ala}**	-	-	-	-	-
***Nme*DAH7PS^{Gln172Ala}**	30 ± 4	97 ± 8	47.5 ± 2.1	1.6	0.5
***Nme*DAH7PS^{Glu176Ala}**	17 ± 2	45 ± 3	15.3 ± 0.6	0.9	0.3
***Nme*DAH7PS^{Glu176Gln}**	15 ± 1	79 ± 8	23.5 ± 0.8	1.6	0.3
***Nme*DAH7PS^{WT}**	15 ± 1	37 ± 2	27.1 ± 0.2	1.8	0.7

Table 5.3.1.1 Kinetic parameters of NmeDAH7PS^{Glu98Ala}, NmeDAH7PS^{Lys107Ala}, NmeDAH7PS^{Gln172Ala}, NmeDAH7PS^{Glu176Ala} and NmeDAH7PS^{Glu176Gln}. All kinetic assays were carried out at high concentrations of other substrate (PEP-100 μM and E4P 250–500 μM).

Kinetic parameters for the *Nme*DAH7PS^{Lys107Ala} variant enzyme could not be determined as this enzyme was found to be completely inactive during testing. The K_M for substrates, E4P and PEP in the *Nme*DAH7PS^{Glu98Ala} variant are 29 ± 2 μM and 2 ± 0.2 μM respectively, which indicate an increase in affinity for PEP compared to *Nme*DAH7PS^{WT}. However, the k_{cat} of this variant is also less than half of the *Nme*DAH7PS^{WT}.

Interestingly, the *Nme*DAH7PS^{Gln172Ala} variant enzyme exhibits increased kinetic parameters for both substrates, with K_M values of 97 ± 8 µM and 30 ± 4 µM for E4P and PEP respectively. The k_{cat} for this variant enzyme however has increased to 47.5 ± 2.1 s^{-1}. The intimate association between this residue and residues, Lys99 and Arg101 of the conserved KPR(T/S) motif on the $\beta_2\alpha_2$ loop observed in the PEP-bound *Nme*DAH7PS^{WT} structure, may have interesting implications for catalysis in this variant.

The K_M of substrates, E4P and PEP for the *Nme*DAH7PS^{Glu176Ala} variant are 45 ± 3 µM and 17 ± 2 µM respectively, which are similar to that of *Nme*DAH7PS^{WT}, however, the k_{cat} of 15.3 ± 0.6 s^{-1} is less than that of the *Nme*DAH7PS^{WT}. The *Nme*DAH7PS^{Glu176Gln} variant has a K_M for PEP of 15 ± 2 µM, which is similar to the *Nme*DAH7PS^{WT}. The K_M for E4P was 79 ± 9 µM and it remains unclear as to why the affinity for E4P is decreased. The k_{cat} for *Nme*DAH7PS^{Glu176Gln} of 23.5 ± 0.8 s^{-1} is also similar to that of the *Nme*DAH7PS^{WT}.

Inhibition kinetic assays were carried out on all variant enzymes in the presence of 50–1000 µM Phe as previously described (section 7.1.27) and the results are displayed in figure 5.3.1.1.

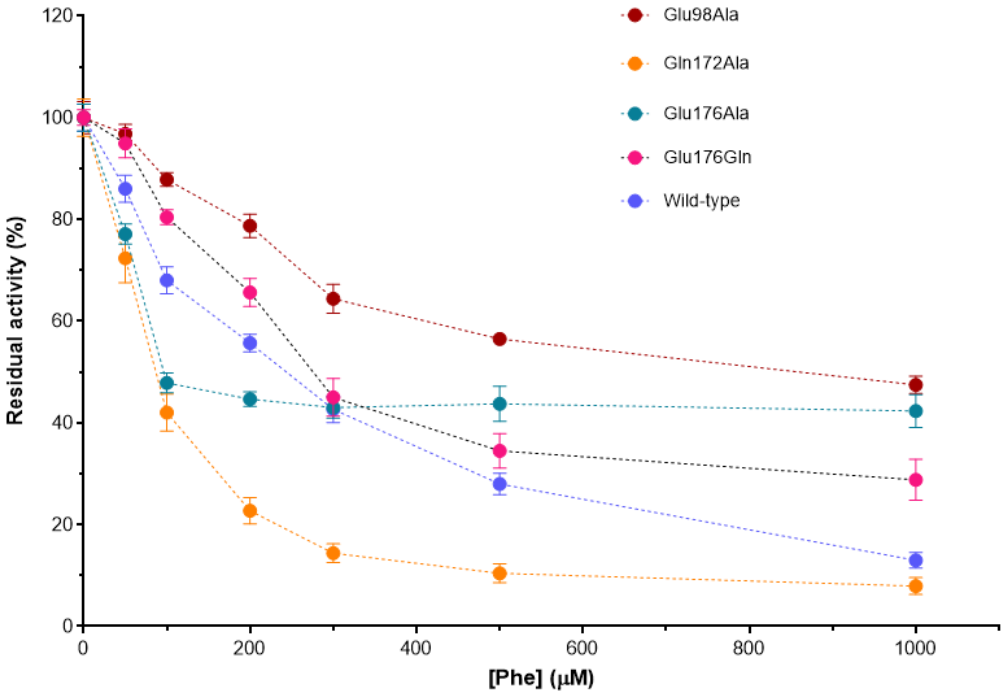


Figure 5.3.1.1. Inhibition plot for all variant enzymes at varying phenylalanine concentrations (50-1000 µM). NmeDAH7PS^{Glu98Ala} variant (brown), NmeDAH7PS^{Gln172Ala} variant (orange), NmeDAH7PS^{Glu176Ala} variant (teal),

NmeDAH7PS^{Glu176Gln} variant (pink) and NmeDAH7PS^{WT} (blue). Saturating substrate concentrations were used for all measurements.

The allosteric regulation of *Nme*DAH7PS^{Glu98Ala} was severely attenuated at all concentrations of Phe tested, indicating that this residue is somehow involved in the communication of allosteric signal. Conversely, *Nme*DAH7PS^{Gln172Ala} showed increased sensitivity to Phe. Interestingly, the inhibitory response of *Nme*DAH7PS^{Glu176Ala} at low concentrations of Phe is increased, but at high concentrations (> 200 µM), the effect plateaus, implying that there is a limit to the regulatory effect of Phe in the *Nme*DAH7PS^{Glu176Ala} variant. This residue may contribute to the formation of the Phe-bound state of the enzyme. A similar mutation involving the same residue in *Nme*DAH7PS^{Glu176Gln} showed a similar trend in sensitivity to Phe as *Nme*DAH7PS^{WT} at all concentrations tested. The response however, was somewhat attenuated.

5.3.2 Hydrogen bonding network variants adopt tetrameric solution state quaternary structure

Analytical SEC was employed to determine the solution-state quaternary structure of the aforementioned variant enzymes. All variant enzymes were shown to elute after the same volume of isocratic flow as *Nme*DAH7PS^{WT}, demonstrating retention of the native tetrameric assembly in solution (Figure 5.3.2.1).

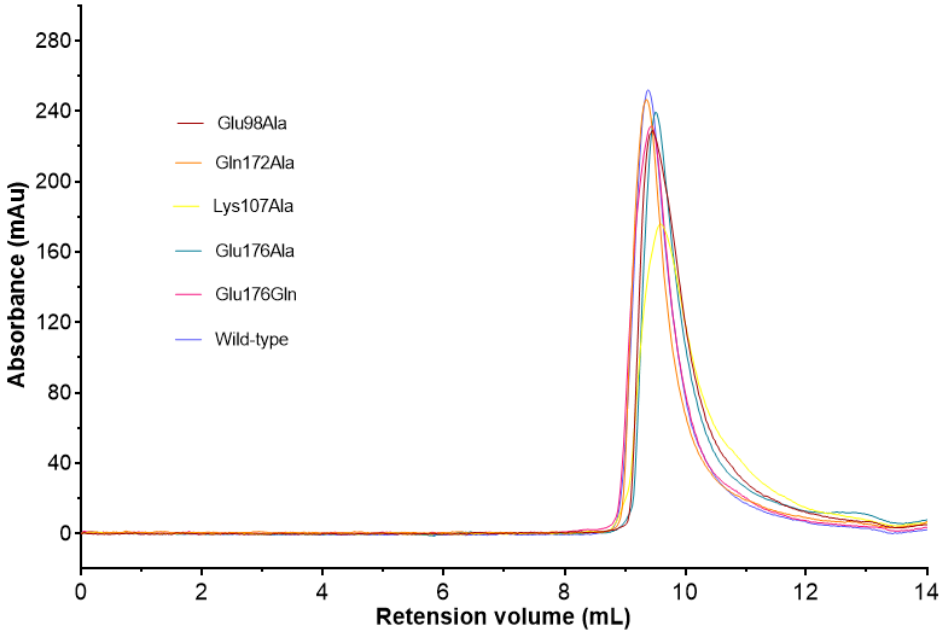


Figure 5.3.2.1 Overlaid Analytical size exclusion chromatography traces of NmeDAH7PS^{WT} (blue), NmeDAH7PS^{Glu98Ala} (brown), NmeDAH7PS^{Lys107Ala} (yellow), NmeDAH7PS^{Gln172Ala} (orange), NmeDAH7PS^{Glu176Ala} (light blue), NmeDAH7PS^{Glu176Gln} (pink). Concentrations used for the experiment varied from 0.6–0.7 mg/mL.

5.3.3 Phe binding characteristics of hydrogen bonding network variants

The total free energy (Δ*G*) of binding, which can be determined experimentally using ITC, is multifactorial and represents the contribution from a number of molecular interactions including: van der Waals interactions, hydrogen bonding, hydration, conformational entropy, electrostatic and ionisation effects, contributions due to conformational transitions and the loss of translational and rotational freedom.[136] The enthalpic and entropic contributions to Δ*G* can be obtained from these isotherms via the Gibbs free energy equation (Δ*G* = Δ*H* - TΔ*S*). This provides interesting insight into the intrinsic properties of Phe binding when compared across a range of variant enzymes.

The dissociation constants and thermodynamics of Phe binding to the aforementioned variant enzymes were determined by ITC. Phe was titrated into the calorimeter cell containing *Nme*DAH7PS variants and binding isotherms obtained. The binding isotherm of Phe to *Nme*DAH7PS^{Glu98Ala}, *Nme*DAH7PS^{Gln172Ala} and *Nme*DAH7PS^{Glu176Gln} were fit to a two-site sequential binding model, like that of the *Nme*DAH7PS^{WT}. However, the binding isotherm of *Nme*DAH7PS^{Glu176Ala} and *Nme*DAH7PS^{Lys107Ala} could only be fit to a one-site model, indicating a significant change in the characteristics of Phe binding, especially with respect to the cooperativity between allosteric sites of the enzyme.

The K_D for Phe binding to each site of *Nme*DAH7PS^{Glu98Ala} was 5 ± 1 µM and 29 ± 5 µM respectively, illustrating the reduction in affinity for both sites compared to *Nme*DAH7PS^{WT} (Figure 5.3.3.1A). From the thermodynamic parameters of *Nme*DAH7PS^{Glu98Ala} it can be inferred that binding at the first site is entropically driven, as is the case for *Nme*DAH7PS^{WT}. The contribution to the free energy of binding at the second site however consists more equally of entropic and enthalpic components compared to *Nme*DAH7PS^{WT}. The variation in binding characteristics for the second site, may explain the reduction in binding affinity observed. Overall the binding is slightly more entropically favourable than *Nme*DAH7PS^{WT} in the *Nme*DAH7PS^{Glu98Ala} variant.

The K_D values for the binding of Phe at both sites of *Nme*DAH7PS^{Gln172Ala} were 14 ± 4 µM and 17 ± 3 µM respectively (Figure 5.3.3.1B). This indicates weaker binding at the first site than *Nme*DAH7PS^{WT}, whereas binding at the second site remains similar to *Nme*DAH7PS^{WT}. This may suggest a small disruption in the cooperativity between allosteric sites. The thermodynamic parameters obtained from the *Nme*DAH7PS^{Glu172Ala} binding isotherm indicate that binding at both sites remains similar to *Nme*DAH7PS^{WT}. However, binding at the second site is slightly more entropically unfavourable and this may contribute to the similarity in the binding affinity of Phe at both sites.

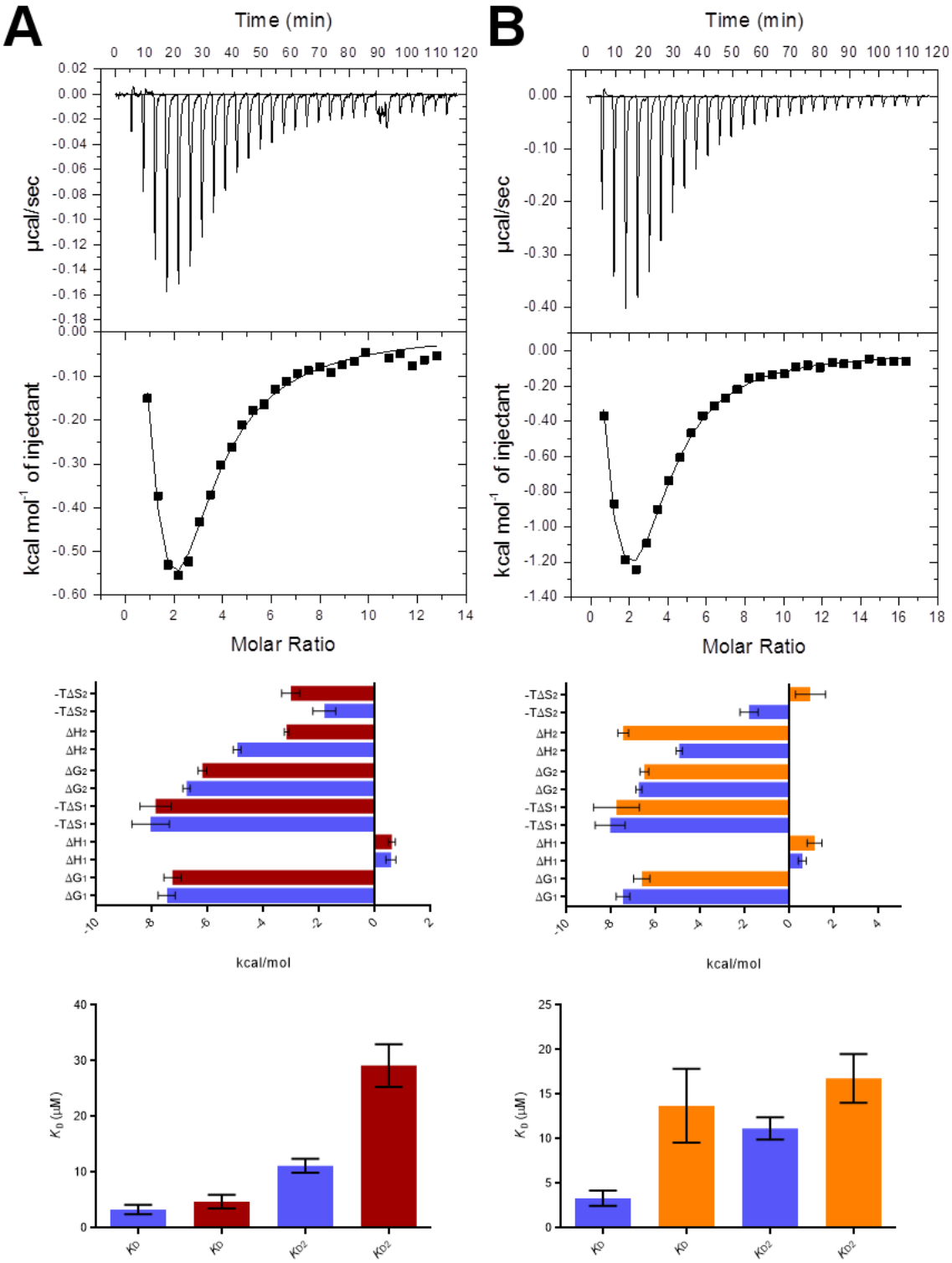


Figure 5.3.3.1 A). ITC data for the titration of 1.2 mM Phe into 18 µM NmeDAH7PS^{Glu98Ala} variant (brown) B) ITC data for the titration of 1.2 mM Phe into 16 µM NmeDAH7PS^{Gln172Ala} variant (orange) Top: Raw data from ITC experiment fit to the two-site sequential binding model Middle. Comparison of the thermodynamic parameters for binding at both sites with NmeDAH7PS^{WT} (blue) Bottom. Comparison of dissociation constant (K_D) at both sites of the sequential binding model with NmeDAH7PS^{WT} (blue). All ITC isotherms were fit and their corresponding heat of dilutions subtracted using OriginLab® version 7.0.

The K_D of Phe binding to the inactive *Nme*DAH7PS^{Lys107Ala} variant, obtained from a one site binding model is 1.0 ± 0.1 μM, indicative of tighter binding than the *Nme*DAH7PS^{WT} (Figure 5.3.3.2A). The molar ratio for Phe binding to this variant was only 0.47 after 48 hours of dialysis. Crystallographic studies have shown that Phe was present in the crystal structure at ~0.6-0.7 occupancy when it was excluded from the crystallisation condition (section 7.6.5). This is consistent with the molar ratio of ligand to enzyme observed by ITC before dialysis over 48 hours, giving a small increase in molar ratio for the titration (Figure A.5.4). This may explain why only a single site binding model could be fit to the data. Unfortunately attempts to dialyse Phe totally out of the enzyme before titration were unsuccessful.

The low K_D for this variant, when Phe was occupying ~half of all binding sites is intriguing. The binding isotherm indicates a very large enthalpic contribution to binding at the allosteric site as well as a highly unfavourable entropic contribution. The presence of Phe in this variant implies that comparisons should only be made with the second binding event of *Nme*DAH7PS^{WT}. Interestingly, the entropic component to binding in of Phe to *Nme*DAH7PS^{Lys107Ala} implies that binding is now driven primarily by electrostatic interactions (ΔH contributions) and that the entropic penalty incurred upon Phe binding has already been paid. Unfortunately, the thermodynamic parameters obtained from this binding isotherm may be influenced by the reduction in binding sites accessible to Phe and therefore deductions based on these characteristics remain tentative.

The K_D of Phe binding to the *Nme*DAH7PS^{Glu176Ala} variant, obtained from a one site binding model was 1.5 ± 0.1 μM (Figure 5.3.3.2B) with a molar ratio of 0.9. Inspection of the thermodynamic parameters of this binding interaction shows that there is a large enthalpic contribution to binding of Phe at this site, with no entropic contribution. The fit of this binding isotherm to a one site binding model implies that the un-cooperative allostery between the Phe binding sites exhibited by *Nme*DAH7PS^{WT} is lost in this variant.

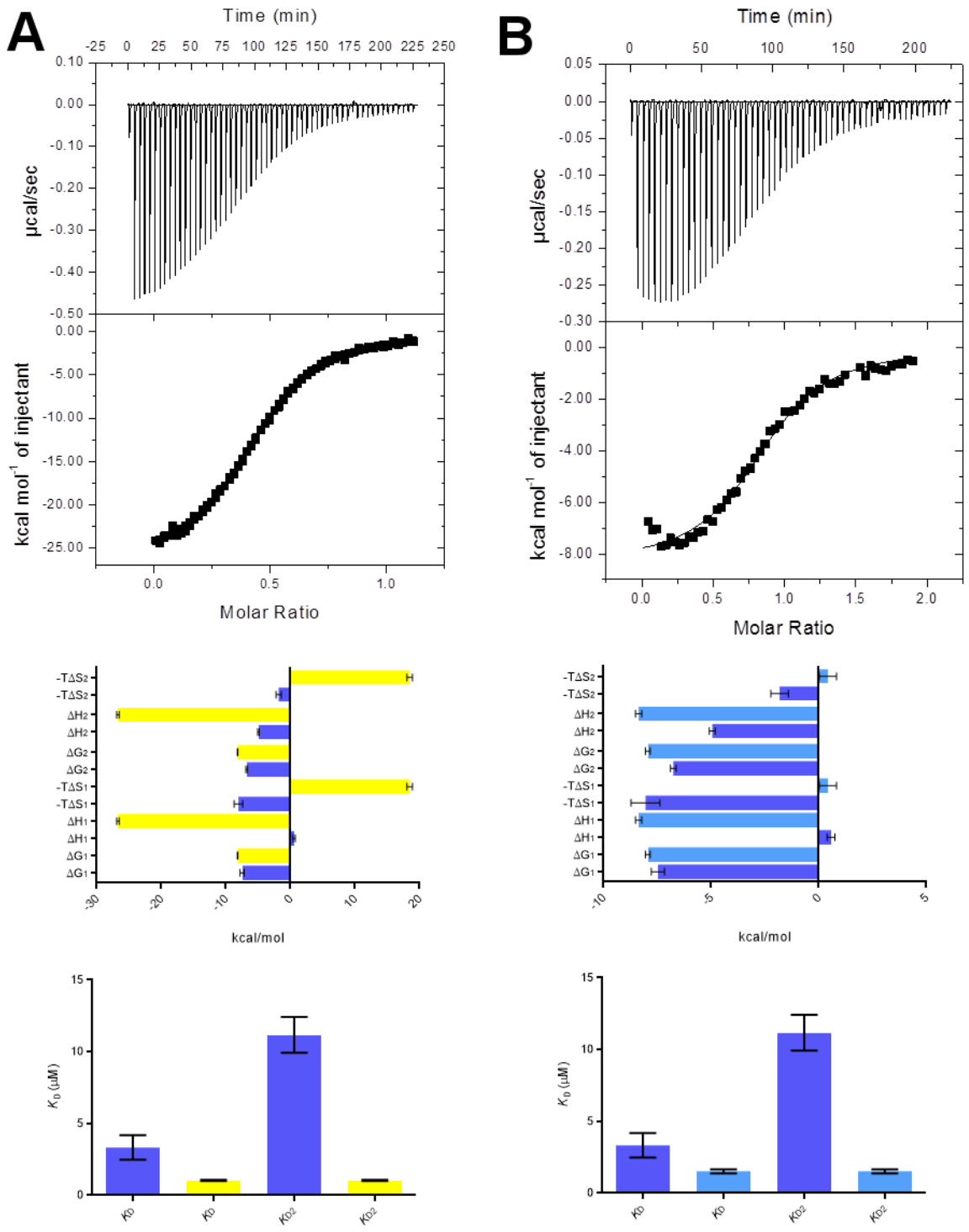


Figure 5.3.3.2 A). ITC data for the titration of 0.11 mM Phe into 21µM NmeDAH7PS^{Lys107Ala} variant (yellow) B) ITC data for the titration of 0.2 mM Phe into 22 µM NmeDAH7PS^{Glu176Ala} variant (light blue) Top. Raw data from ITC experiment fit to a one-site binding model. Middle. Comparison of the thermodynamic parameters for binding at both sites with NmeDAH7PS^{WT} (blue). Bottom. Comparison of dissociation constant (K_D) at both sites of the sequential binding model with NmeDAH7PS^{WT} (blue). All ITC isotherms and graphs were fit and their corresponding heat of dilutions subtracted using OriginLab® version 7.0.

Phe binding is tighter in these two variant enzymes (*Nme*DAH7PS^{Lys107Ala} and *Nme*DAH7PS^{Glu176Ala}) than *Nme*DAH7PS^{WT}. The thermodynamic characteristics are also similar, and thus implicate that the salt-bridge formed between these two residues in the PEP-bound *Nme*DAH7PS^{WT} structure is important for the cooperativity between binding sites during Phe binding. The vastly different thermodynamic character for Phe binding to these variants indicates that in the absence of this salt bridge, the entropic cost for Phe binding may already be satisfied before titration of Phe into the calorimeter.

The K_D values for Phe binding at both sites of *Nme*DAH7PS^{Glu176Gln} were 2 ± 1 µM and 9 ± 1 µM respectively, and are similar to that of *Nme*DAH7PS^{WT} (Figure 5.3.3.3). This is perhaps not surprising, as the interactions that Gln176 makes with surrounding residues are likely to be similar to those made by Glu176 in *Nme*DAH7PS^{WT}. Inspection of the thermodynamic parameters illustrate that binding of Phe to this variant has a greater enthalpic contribution to binding at both sites than *Nme*DAH7PS^{WT}.

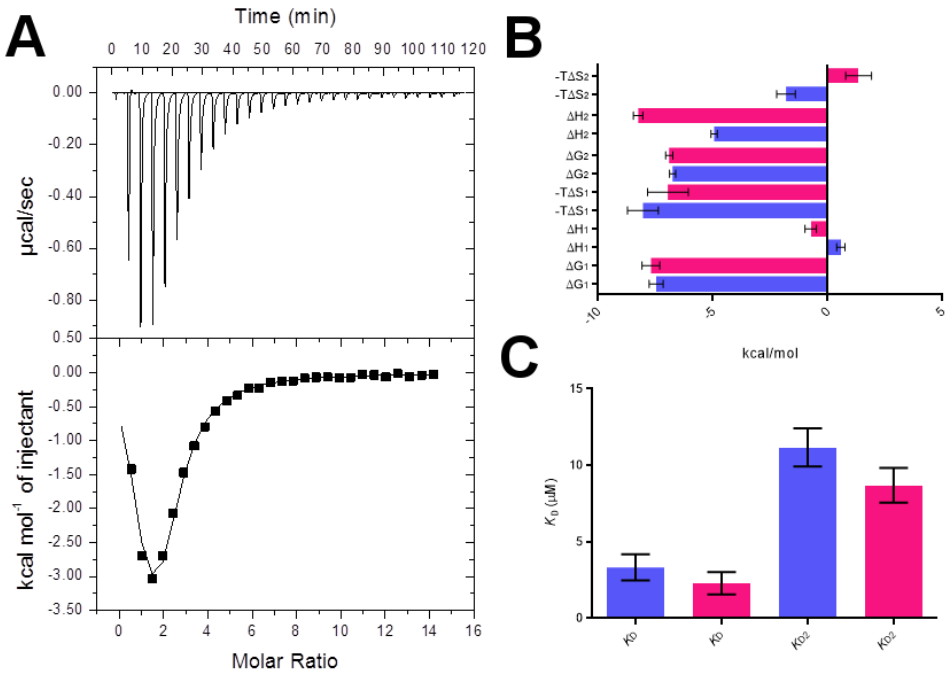


Figure 5.3.3.3 ITC data for the titration of 1.2 mM Phe into 16 µM NmeDAH7PS^{Glu176Gln} variant (pink) A) Raw data from ITC experiment fit to two-site sequential binding model B) Comparison of the thermodynamic parameters for binding at both sites with NmeDAH7PS^{WT} (blue) C) Comparison of dissociation constant (K_D) at both sites of the two-site sequential binding model with NmeDAH7PS^{WT} (blue). All ITC isotherms and graphs were fit and their corresponding heat of dilutions subtracted using OriginLab® version 7.0.

5.3.4 Crystallographic characterisation of hydrogen bonding network variants

Crystal structures of these variant enzymes in the presence and absence of Phe were obtained using X-ray diffraction methods, in order to structurally rationalise the characteristics of Phe binding and the kinetic properties observed for these variant enzymes. Crystals of the *Nme*DAH7PS variants were grown using procedures describe in sections 7.1.30 and 7.1.31. The soaking of Phe into these structures was achieved using the procedures described in section 7.1.33.

The structural differences in these variants, not only compared to the *Nme*DAH7PS^{WT}, but to other variant structures in the presence and absence of Phe, may illuminate the importance of interactions between the mutated residues in the network. The structural variation upon removal of the ionisable side-chains of residues involved in the network may better our understanding of how allostery is communicated between distant sites of *Nme*DAH7PS and help illustrate the role of each residue in the transfer of this allosteric signal. All ligands were present at full occupancy unless otherwise stated.

Comparison of the RMSD per residue for the tight dimer of all Phe-absent variant enzymes to the PEP-bound *Nme*DAH7PS^{WT} structure (PDB code: 4HSN), indicate that the region with the most significant structural variation is the $\beta_2\alpha_2$ loop (res. 95-120). Comparisons of the RMSD per residue for the tight dimer of the Phe-bound variant structures with the Phe-bound *Nme*DAH7PS^{WT} (PDB code: 4UC5) indicate that the regions of most significant variation is also localised to the $\beta_2\alpha_2$ loop (res 95-120) as well as the $\beta_8\alpha_8$ loop and α_8 helix (res. 305-349).

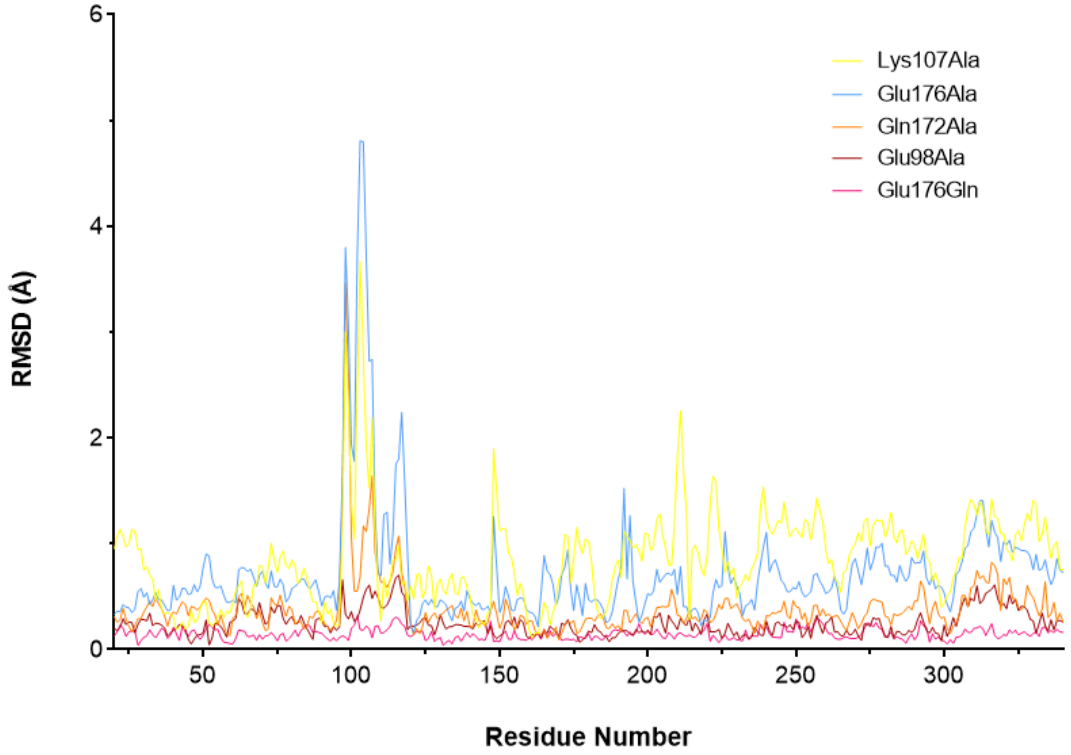


Figure 5.3.4.1 Comparison of the root mean squared deviation of all Phe-absent variant enzymes against the PEP-bound NmeDAH7PS^{WT} structure (PDB code: 4HSN) NmeDAH7PS^{Glu98Ala} (brown), NmeDAH7PS^{Gln172Ala} (orange), NmeDAH7PS^{Lys107Ala} (yellow), NmeDAH7PS^{Glu176Ala} (light blue), NmeDAH7PS^{Glu176Gln} (pink).

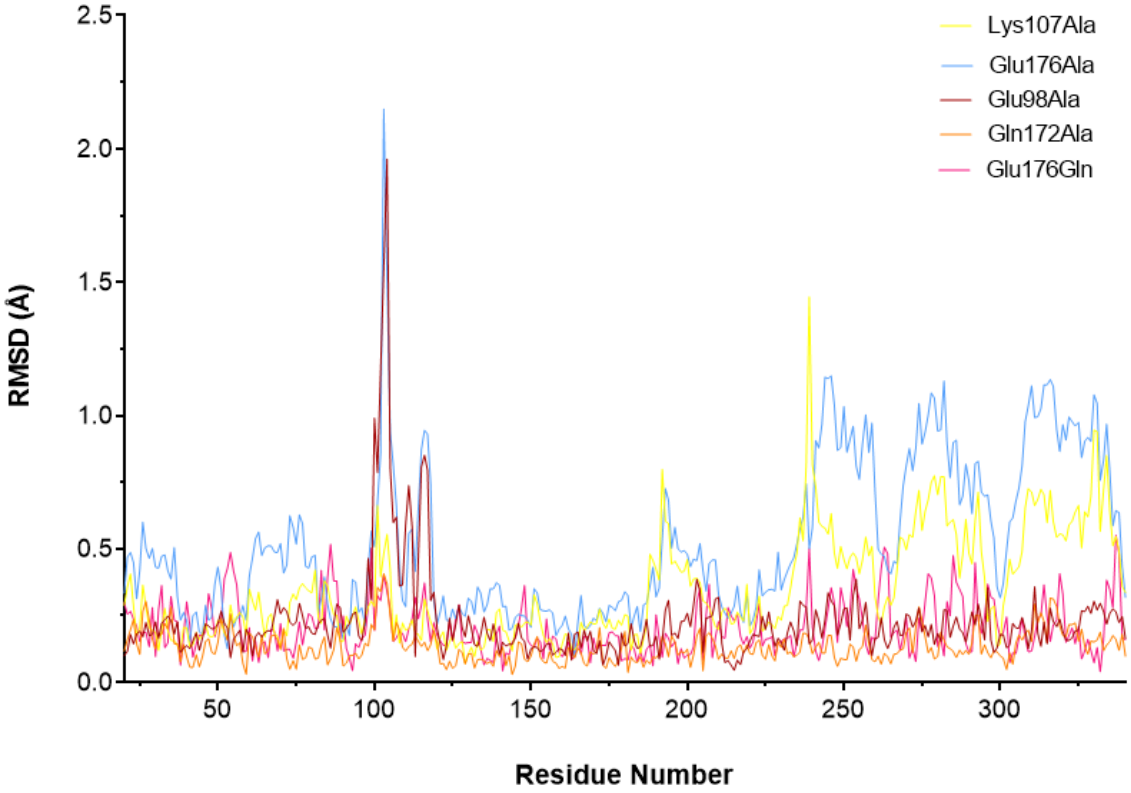


Figure 5.3.4.2 Root mean squared deviation of all Phe-bound variant enzymes against Phe-bound NmeDAH7PS^{WT} structure PDB code: 4UC5. NmeDAH7PS^{Glu98Ala} (brown), NmeDAH7PS^{Lys107Ala} (yellow), NmeDAH7PS^{Gln172Ala} (orange), NmeDAH7PS^{Glu176Ala} (light blue) and NmeDAH7PS^{Glu176Gln} (pink).

Crystal structures of Phe- and PEP-bound NmeDAH7PS^{Glu98Ala}

Two crystal structures of the *Nme*DAH7PS^{Glu98Ala} variant were determined by X-ray diffraction. The PEP-bound structure was refined to a resolution of 2.50 Å and crystallised in the same space group as the PEP-bound *Nme*DAH7PS^{WT} structure ($P12_11$) (PDB code: 4HSN). Data collection and refinement statistics are shown in Table 5.3.4.1. The RMSD for the alignment of the tight dimers of the PEP-bound *Nme*DAH7PS^{Glu98Ala} variant to the PEP-bound *Nme*DAH7PS^{WT} structure was 0.235-0.358 Å. This indicates a very similar overall structure for this variant enzyme. These structures are remarkably similar (Figure 5.3.4.1), however the most drastic difference between the two structures is localised to the $\beta_2\alpha_2$ loop on which this mutation sits.

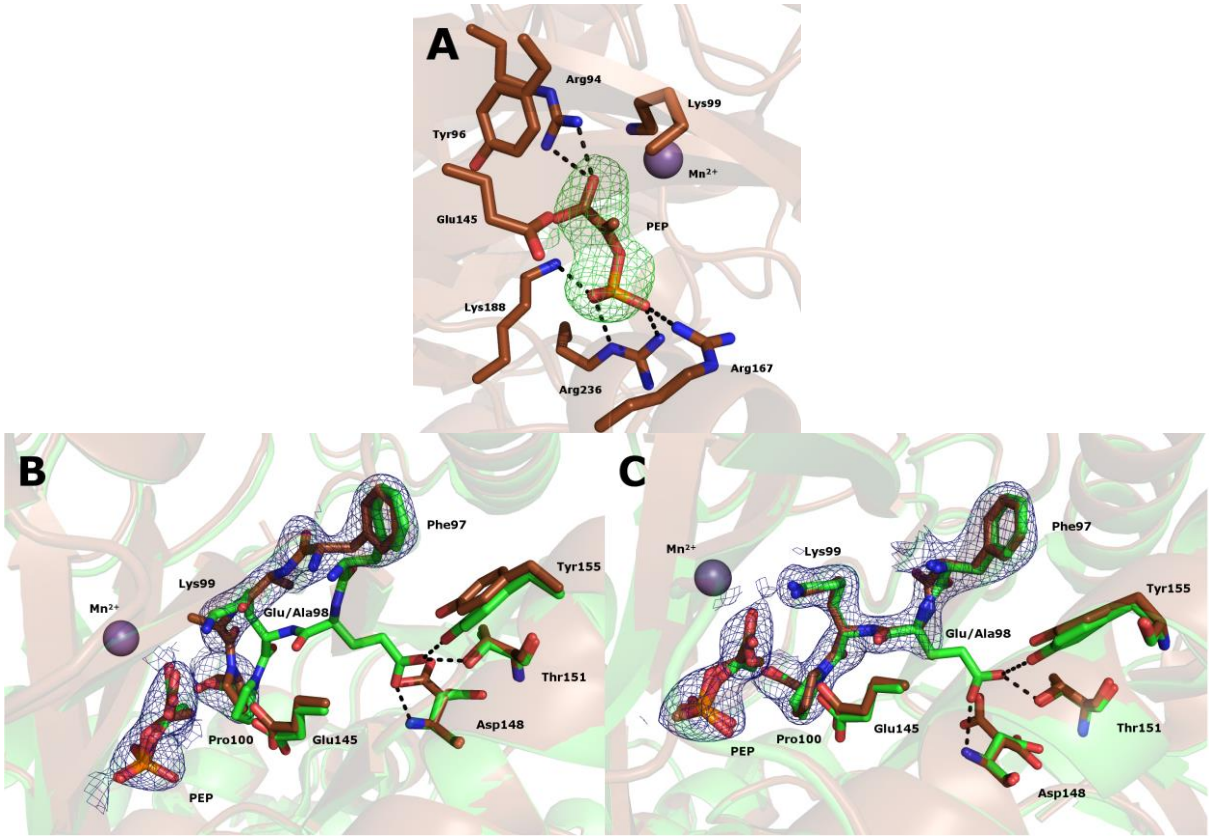


Figure 5.3.4.3 A) $|F_o\text{-}F_c|$ map coloured in green and contoured at 3.0 σ to illustrate the presence of PEP at full occupancy in the active site. B) Structural alignment of PEP-bound NmeDAH7PS^{Glu98Ala} (brown) (PDB code: 5CZ0) with PEP-bound NmeDAH7PS^{WT} (green) (PDB code: 4HSN) indicating the variability within the same structure upon this mutation and the loss of connection to residues of the $\beta_3\alpha_3$ loop. $|2F_o\text{-}F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate mutation. Chain A, indicating new position of the $\beta_2\alpha_2$ loop C) Chain B, indicating the preservation of the position of the $\beta_2\alpha_2$ loop.

In *Nme*DAH7PS^{Glu98Ala} the hydrogen bonds that link the $\beta_2\alpha_2$ loop with residues of the $\beta_3\alpha_3$ loop and the α_3 helix are lost. This indicates the removal of a crucial link between the allosteric and active site within the same chain, which may contribute to the higher affinity of this variant for

PEP, as PEP binding is no longer mediated by the interaction between the $\beta_2\alpha_2$ loop and residues, Asp148, Thr151, and Tyr155. The variability in conformation of the N-terminal end of the $\beta_2\alpha_2$ loop between chains of the aligned dimer indicates greater flexibility in this region for *Nme*DAH7PS^{Glu98Ala} (Figure 5.3.4.1 B & C). Interestingly, the active site ligand PEP adopts the same conformation to that observed in the PEP-bound *Nme*DAH7PS^{WT} even though the residues around it in at least one of the chains have moved considerably (Figure 5.3.4.3). There is a small repositioning of residue, Glu145 (also located on the $\beta_3\alpha_3$ loop), which is involved in the binding and positioning of the nucleophilic water molecule, responsible for the attack of the oxocarbenium ion reaction intermediate.[100] This residue has moved closer to the C_2 position of PEP, probably due to the loss of interaction between Glu98 and the $\beta_3\alpha_3$ loop and α_3 helix. The restriction on the position of Glu145 incurred by these hydrogen bonds is lost in this mutation and thus, may explain the decreased K_M observed in kinetic assays for PEP in *Nme*DAH7PS^{Glu98Ala}.

The Phe-bound structure of *Nme*DAH7PS^{Glu98Ala} was refined to a resolution of 2.42 Å and crystallised in the same space group as the Phe-bound *Nme*DAH7PS^{WT} structure ($P12_11$) (PDB code: 4UC5). Data collection and refinement statistics are shown in table 5.3.4.1. The RMSD for the alignment of the tight dimers of the Phe-bound *Nme*DAH7PS^{Glu98Ala} variant to the Phe-bound *Nme*DAH7PS^{WT} structure was 0.192-0.381 Å, indicative of a high similarity between these structures (Figure 5.3.4.2). Comparison of these structures indicates that the only region of the protein that adopts a different conformation in the presence of Phe is the $\beta_2\alpha_2$ loop.

The binding of Phe in the allosteric site of the *Nme*DAH7PS^{Glu98Ala} variant enzyme is very similar to the Phe-bound *Nme*DAH7PS^{WT}. This is perhaps not surprising as the distance between Cα of the mutated residue and Cα of allosteric Phe is ~22 Å from the allosteric site of the adjacent chain across the tight dimer interface and ~19 Å from the allosteric site within the same chain. The reorganisation of the network observed in the Phe-bound *Nme*DAH7PS^{WT} cannot occur in *Nme*DAH7PS^{Glu98Ala} variant upon removal of the Glu side-chain. As a result, both Gln172′ and Asn111, which were involved in hydrogen bonds with Glu98 in the Phe-bound *Nme*DAH7PS^{WT} structure adopt different rotamers. The amide functionality of Asn111 is no longer able to form a hydrogen bond with any other residue and Gln172′ does not form hydrogen bonds with residues of the $\beta_2\alpha_2$ loop, but instead forms a hydrogen bond with the sidechain of Glu176′

located on the same helix. The side-chain of Lys107 could not be resolved due to partial disorder as it too is intimately associated with the aforementioned residues in the Phe-bound *Nme*DAH7PS^{WT} structure. This may be an artefact of the resolution that this structure was determined at, as the side-chains of several residues of the $\beta_2\alpha_2$ loop were also unresolved. The crucial link between the active and allosteric site of the same chain and its influence on communication with the $\beta_2\alpha_2$ loop across the tight dimer interface that this residue affords in *Nme*DAH7PS^{WT} may explain the reduced sensitivity to Phe observed in kinetic studies for this variant.

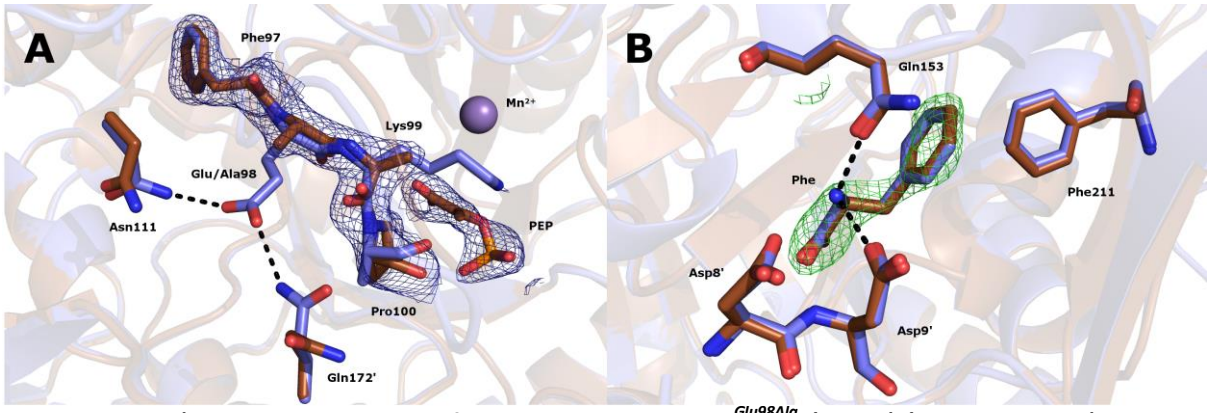


Figure 5.3.4.4 A) Structural alignment of Phe-bound NmeDAH7PS^{Glu98Ala} (brown) (PDB code: 5CZS) with Phe-bound NmeDAH7PS^{WT} (blue) (PDB code: 4UC5) indicating the inability of res. 98 to form the hydrogen bonds that it does in the NmeDAH7PS^{WT}. $|2F_o\text{-}F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate mutation, the similar positioning of this loop and the presence of PEP (0.75 occ.) in its upright conformation at the active site of NmeDAH7PS^{Glu98Ala} B) Allosteric binding site indicating the conservation of structural characteristics in NmeDAH7PS^{Glu98Ala}. $|F_o\text{-}F_c|$ map coloured in green and contoured at 3.0 σ to illustrate presence of Phe at full occupancy.

Interestingly, the presence of PEP is evident at the active site of the Phe-bound *Nme*DAH7PS^{Glu98Ala} variant enzyme and has been modelled in at 0.75 occupancy. Chapter 4 showed that in the presence of Phe, PEP adopted a flipped binding mode. However, in the *Nme*DAH7PS^{Glu98Ala} variant enzyme, PEP binds in its normal reactive position (albeit at 0.75 occupancy) in a conformation that resembles that observed in the active enzyme. The interactions that this residue makes with other residues of the $\beta_2\alpha_2$ loop, as well as Gln172′ and Glu176′ in the Phe+PEP-bound *Nme*DAH7PS^{WT} structure (PDB code: 5DCB) are therefore likely to be important to the flipped binding mode of PEP. In the absence of the Glu side-chain, the active site is not sufficiently reorganised to bind PEP in its flipped binding mode. This may be a significant contributor to the reduction in allosteric response observed from kinetic studies in *Nme*DAH7PS^{Glu98Ala} (Figure 5.3.1.2).

Crystal structures of Phe- and PEP-bound NmeDAH7PS^{Gln172Ala}

Two crystal structures of the *Nme*DAH7PS^{Gln172Ala} variant were determined by X-ray diffraction. The PEP-bound structure was refined to a resolution of 2.68 Å and crystallised in the same space group as *Nme*DAH7PS^{WT} ($P12_11$) (PDB code: 4HSN). Data collection and refinement statistics are shown in table 5.3.4.1. The RMSD for the alignment of the tight dimers of the PEP-bound *Nme*DAH7PS^{Gln172Ala} variant to the *Nme*DAH7PS^{WT} structure (PDB code: 4HSN) was 0.270-0.449 Å. These structures are similar, but again show greater variation, localised to the $\beta_2\alpha_2$ loop (Figure 5.3.4.1).

Gln172′ in the PEP-bound *Nme*DAH7PS^{WT} structure forms an important hydrogen bond with the backbone carbonyl and peptide amino group of Lys99 and Arg101 respectively. The importance of Lys99 to catalysis[100] and its positioning at the N-terminal end of the $\beta_2\alpha_2$ loop highlight the possible influence of Gln172′ on catalysis. In the *Nme*DAH7PS^{Gln172Ala} variant enzyme, the $\beta_2\alpha_2$ loop from the only chain that could be resolved was shifted with an average RMSD from the alignment of the tight-dimer of 1.42 Å. The partial disorder of the $\beta_2\alpha_2$ loop in three of the four chains of the tetramer is indicative of increased flexibility in this region and may be a result of the loss of the aforementioned interactions with the $\beta_2\alpha_2$ loop. In the chain where the $\beta_2\alpha_2$ loop could be resolved, interactions between residue 172′ and the backbone carbonyl of Lys99, as well as the backbone peptide amino group of Arg101 were lost. The salt bridge between the side-chains of Lys107 and Glu176′ has also been compromised in this variant (Figure 5.3.4.5B). Interestingly, the result of the loss of these interactions are visible in the crystal structure, where Lys107 and Lys99 adopt different positions and conformations in the PEP-bound *Nme*DAH7PS^{Gln172Ala} variant compared to PEP-bound *Nme*DAH7PS^{WT} (Figure 5.3.4.5B).

The partial disorder for residues of the $\beta_2\alpha_2$ loop resulted in the inability to resolve the side-chain of Lys107 in all four chains. While this is probably a result of the failure to form a salt bridge with Glu176′, it may be due to the low resolution that this dataset was collected at. The new position of this loop in the presence of PEP may have direct implications for the binding of substrates. Lys99 adopts a significantly different position and is no longer in a conformation able to form a hydrogen bond with the carboxylate of PEP. The new conformation of residues

(97-101) at the N-terminal end of this loop may also contribute the increased K_M for substrates, PEP and E4P. The hydrogen bonds between Glu98 and residues Asp148, Thr151 and Tyr155 of the $\beta_3\alpha_3$ loop and α_3 helix are no longer present in the *Nme*DAH7PS^{Gln172Ala} variant structure. The previously outlined residues involved in the formation of Phe-bound network; Glu98, Lys99 and Lys107 all adopt different conformations in this variant in the presence of PEP compared to the PEP-bound *Nme*DAH7PS^{WT} structure. Therefore, *Nme*DAH7PS^{Gln172Ala} may be primed for inhibition by Phe due to the lack of stabilising effect provided by Ala172′ on the $\beta_2\alpha_2$ loop. This may contribute to the hyper-sensitivity of this variant to concentrations of Phe.

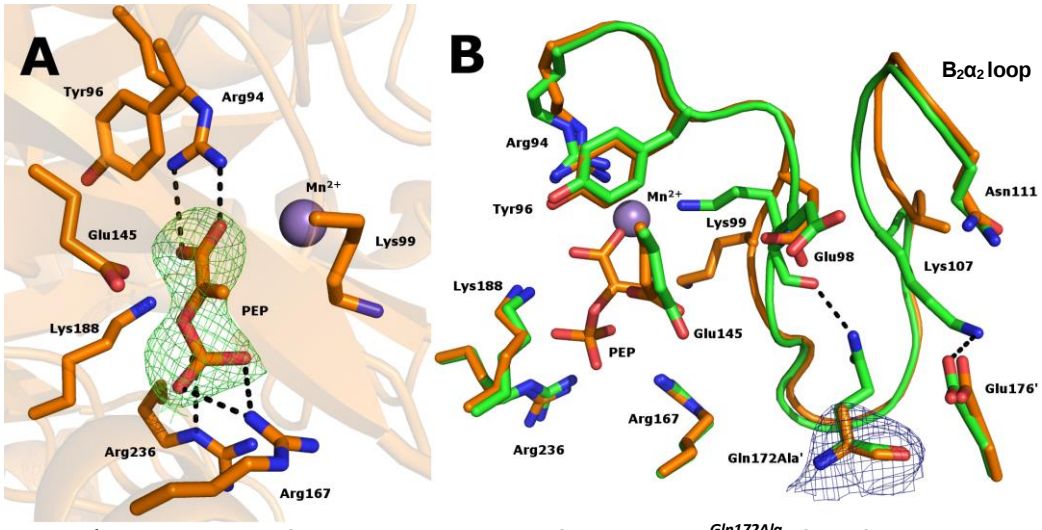


Figure 5.3.4.5 A) Binding mode of PEP in the active site of NmeDAH7PS^{Gln172Ala}. $|F_o\text{-}F_c|$ map coloured in green and contoured at 3.0 σ to illustrate presence of PEP in the site at full occupancy. B) Comparison of the positioning of the $\beta_2\alpha_2$ loop in NmeDAH7PS^{Gln172Ala} (orange) with that of the NmeDAH7PS^{WT} (green) (PDB code: 4HSN) illustrating the loss of hydrogen bonds between residues of the $\beta_2\alpha_2$ loop and Ala172 in the NmeDAH7PS^{Gln172Ala} variant. $|2F_o\text{-}F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate the mutation.

The Phe-bound structure was refined to a resolution of 2.30 Å and crystallised in the same space group as the PEP-bound *Nme*DAH7PS^{WT} ($P12_11$). Data collection and refinement statistics are shown in table 5.3.4.1. The RMSD for the alignment of the tight dimers of the Phe-bound *Nme*DAH7PS^{Gln172Ala} variant to the Phe *Nme*DAH7PS^{WT} structure (PDB code: 4UC5) was 0.140-0.252 Å. This is indicative of a very similar overall structure for this variant enzyme and there are no regions of obvious variation (Figure 5.3.4.2).

The Phe binding site for the *Nme*DAH7PS^{Gln172Ala} remained unaffected by the mutation and the residues that make up the site were in the same position as was observed in the Phe-bound

*Nme*DAH7PS^{WT} structure (Figure 5.3.4.6A). The $\beta_2\alpha_2$ loop also adopts the same conformation as that of the Phe-bound *Nme*DAH7PS^{WT} structure (Figure 5.3.4.6B). This implies that the allosteric effect of Phe binding on this loop and the surrounding residues remains largely unaffected by this mutation and indicates that it may not influence the formation of the Phe-bound network described previously. It must be noted that the conformations explored by the protein in the crystalline form are limited and may not totally reflect what may occur in solution, therefore, the similarity between conformations of the $\beta_2\alpha_2$ loop may be an artefact of crystal packing effects during the soaking experiment. However, the similarity between the Phe-bound *Nme*DAH7PS^{Gln172Ala} structure and the Phe-bound *Nme*DAH7PS^{WT} structure is in accordance with the MD work done in chapter 5, where this residue was not highlighted as being crucial to the formation of the Phe-bound "regulated" network.

Taken together with what was observed in the PEP-bound *Nme*DAH7PS^{Gln172Ala} variant, it seems that the enzyme may be primed for inhibition by Phe, but can still undergo the reorganisation in the presence of Phe to form the Phe-bound network of interactions. This may explain the hypersensitivity of this variant enzyme to Phe even when Phe binding is slightly weaker than the wild-type enzyme, shown by ITC.

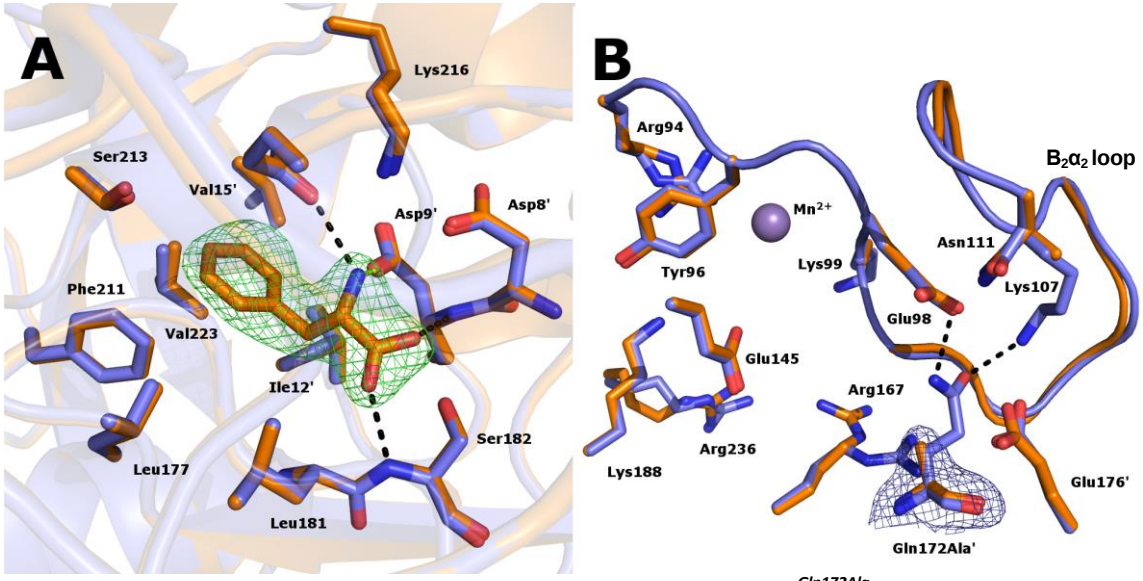


Figure 5.3.4.6 A) Binding mode of Phe in the allosteric site of NmeDAH7PS^{Gln172Ala}. $|F_o-F_c|$ map coloured in green and contoured at 3.0 σ to illustrate presence of Phe in the site at full occupancy. B) Comparison of the positioning of the $\beta_2\alpha_2$ loop in Phe-bound NmeDAH7PS^{Gln172Ala} (orange) with that of the Phe-bound NmeDAH7PS^{WT} (blue) (PDB code: 4UC5), illustrating the loss of hydrogen bonds between residues of the $\beta_2\alpha_2$ loop and Ala172 in the NmeDAH7PS^{Gln172Ala} variant. $|2F_o-F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate the mutation.

Crystal structure of Phe+ PEP-bound NmeDAH7PS^{Lys107Ala}

A single crystal structure of the *Nme*DAH7PS^{Lys107Ala} variant was determined by X-ray diffraction methods. The PEP+Phe double liganded structure was refined to a resolution of 1.75 Å and crystallised in the same space group as the PEP-bound *Nme*DAH7PS^{WT} ($P12_11$) (PDB code: 4HSN). Data collection and refinement statistics are shown in table 5.3.4.2. The RMSD for the alignment of the tight dimers of this variant to the Phe-bound *Nme*DAH7PS^{WT} structure (PDB code: 4UC5) was 0.216-0.323 Å and the RMSD for the alignment of the tight dimers of this variant to the Phe+PEP *Nme*DAH7PS^{WT} structure (PDB code: 5DCB) was 0.206-0.249 Å. This indicates that this enzyme is again similar to both Phe-bound structures of *Nme*DAH7PS^{WT} but does show subtle differences in the structure of the $\beta_2\alpha_2$ loop. Unfortunately, a crystal structure without the allosteric ligand Phe was unable to be obtained, as even when Phe was not soaked into the crystal, the ligand was present at the allosteric site at ~0.6-0.7 occupancy. This may be due to the high affinity that this enzyme displayed for Phe from the ITC experiments, and indicates that Phe may have remained bound to this variant throughout the purification.

The Phe binding site remained largely intact in the *Nme*DAH7PS^{Lys107Ala} variant enzyme. A structural alignment of the Phe binding site indicates large conservation of the structural features that make up the site (Figure 5.3.4.7A). PEP was bound in this structure in a flipped orientation, very similar to that seen for Phe+PEP-bound *Nme*DAH7PS^{WT} (PDB code: 5DCB) (Figure 5.3.4.7.B). Perhaps what is most interesting, is the conformation adopted by the $\beta_2\alpha_2$ loop in the presence of allosteric ligand. In the *Nme*DAH7PS^{Lys107Ala} variant, Arg101, part of the conserved KPR(T/S) motif (believed to be responsible for binding the phosphate of E4P) adopts a new conformation, where it forms a salt bridge with Glu98 like that observed in the Phe-bound *Eco*DAH7PS and Tyr bound *Sce*DAH7PS structures (PDB codes: 1KFL and 1OF6). Arg101 also forms hydrogen bonds with the backbone carbonyl of Lys99 and the side-chain oxygen of Gln172′ (Figure 5.3.4.8A). This new conformation adopted by Arg101 upon mutation of Lys107, causes a small reorganisation of the hydrogen bonding network of the $\beta_2\alpha_2$ loop.

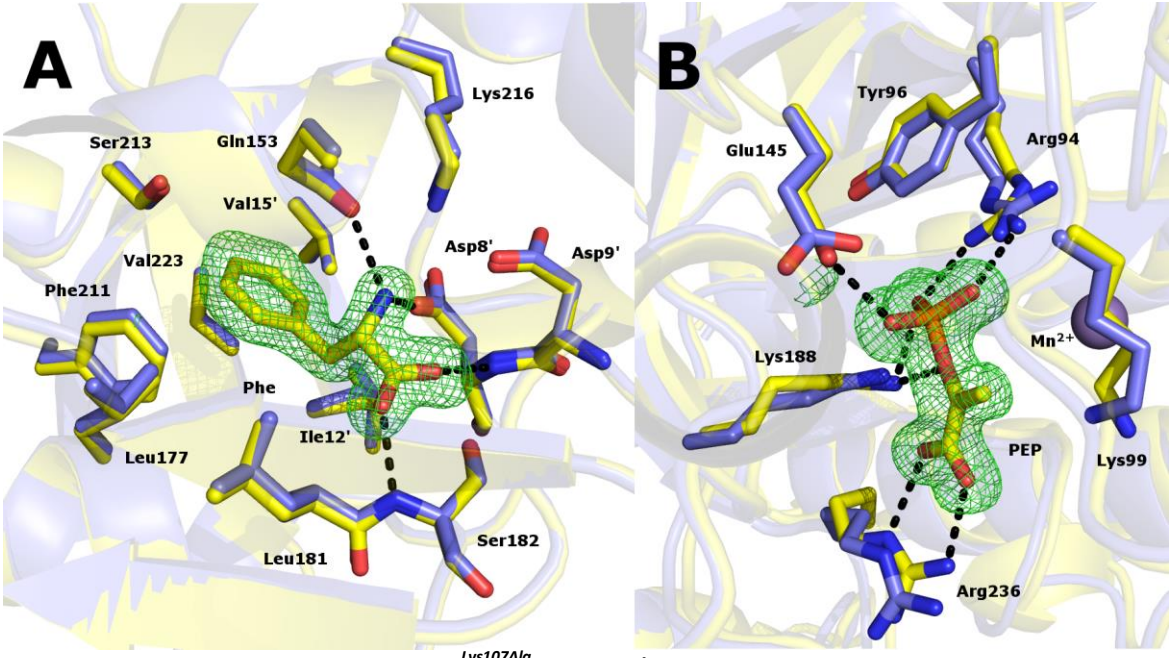


Figure 5.3.4.7 Double liganded NmeDAH7PS^{Lys107Ala} variant A) Allosteric binding site indicating the conservation of structural characteristics in NmeDAH7PS^{Lys107Ala} variant compared to NmeDAH7PS^{WT} (PDB code: 4UC5). $|F_o$-$F_c|$ map coloured in green and contoured at 3.0 σ to illustrate presence of Phe at full occupancy. B) Active site indicating the similarity of the side-chain conformations that make up the site for NmeDAH7PS^{Lys107Ala} (yellow) compared to NmeDAH7PS^{WT} (blue) (PDB code: 4UC5). $|F_o$-$F_c|$ map coloured in green and contoured at 3.0 σ to illustrate presence of PEP in the site, in a flipped binding mode at full occupancy.

Comparison of the new conformation of Arg101 to that observed in the Phe-bound, or Phe+PEP-bound structures of *Nme*DAH7PS^{WT} (PDB codes: 4UC5 and 5DCB) indicates that the ability of this residue to appropriately bind the phosphate functionality of E4P during catalysis may be severely reduced. Interestingly, the conformation of this Arg residue in the double liganded structure from *Eco*DAH7PS (PDB code: 1KFL) is the same as is observed for the *Nme*DAH7PS^{Lys107Ala} variant (Figure 5.3.4.8B). The removal of the Lys107 side-chain may reduce the steric hindrance encountered by Arg101, enabling the formation of a salt bridge with Glu98.

The salt bridge formed between Lys107 and Glu176' across the tight dimer interface in the PEP-bound *Nme*DAH7PS^{WT} structure may have two roles. Firstly, it may be involved in the stabilisation the $\beta_2\alpha_2$ loop during catalysis, ensuring that the KPRT motif is in a conformation appropriate for the binding and catalysis of substrates. Secondly, limiting the ability to form a salt bridge between Arg101 and Glu98, which may result in the tighter, more enthalpically driven binding of Phe, where communication between both Phe binding sites is lost.

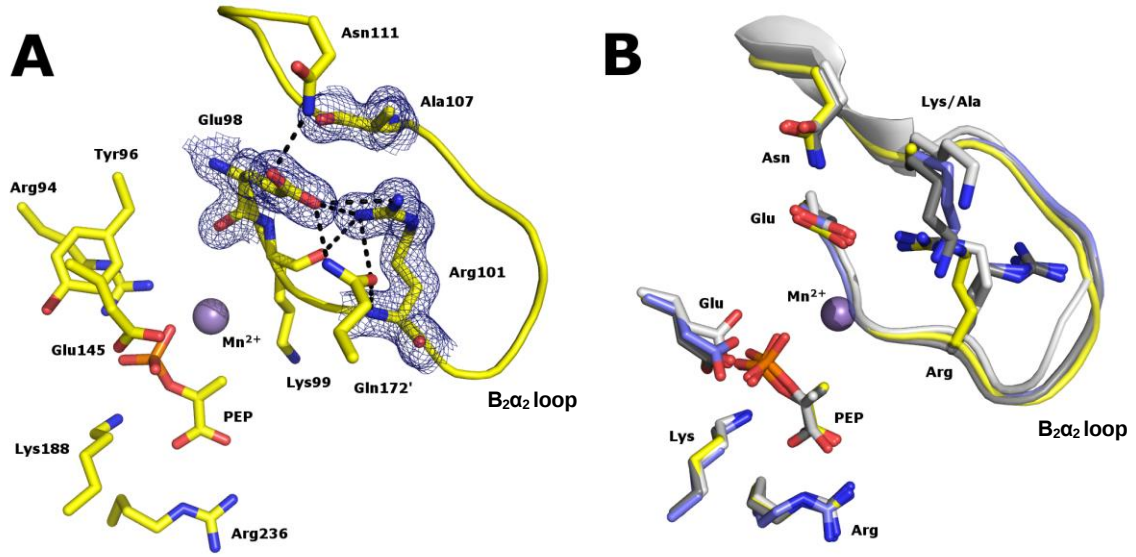


Figure 5.3.4.8 A) NmeDAH7PS^{Lys107Ala} variant active site, illustrating the salt-bridge and hydrogen bonds formed between Arg101 and Glu98, Asn111 and Gln172. $|2F_o\text{-}F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate the new conformation of Arg101 and the mutation of Lys107. B) Active site alignment of double liganded NmeDAH7PS^{Lys107Ala} (yellow) (PDB code: 5D05), Phe-bound NmeDAH7PS^{WT} (blue) (PDB code: 4UC5), Double liganded NmeDAH7PS^{WT} (PDB code: 5DCB) and the double liganded EcoDAH7PS (wihte) (PDB code: 1KFL) indicating the similarity between EcoDAH7PS and the NmeDAH7PS^{Lys107Ala} variant in the conformation adopted by the $\beta_2\alpha_2$ loop.

Crystal structures of Phe-bound and apo NmeDAH7PS^{Glu176Ala}

Two crystal structures of the *Nme*DAH7PS^{Glu176Ala} variant were determined by X-ray diffraction methods. Unfortunately, attempts to obtain a PEP-bound structure of this variant were not successful. However, the apo structure was refined to a resolution of 2.40 Å and crystallised in an orthorhombic space group ($P2_12_12_1$). Data collection and refinement statistics are shown in table 5.3.4.2. The RMSD for the alignment of the tight dimers of the apo *Nme*DAH7PS^{Glu176Ala} variant to the PEP-bound *Nme*DAH7PS^{WT} structure (PDB code: 4HSN) was 0.537-0.710 Å. The RMSD for the alignment of the tight dimers of the apo *Nme*DAH7PS^{Glu176Ala} variant to the Phe-bound *Nme*DAH7PS^{WT} structure (PDB code: 4UC5) was 0.862-1.017 Å. This indicates that there are differences between the apo *Nme*DAH7PS^{Glu176Ala} enzyme and both the Phe- and PEP-bound *Nme*DAH7PS^{WT} structures. Analysis of this RMSD per residue indicates again that the region with the most significant variation is the $\beta_2\alpha_2$ loop (Figure 5.3.4.1).

Glu176′ is intrinsically involved in the stabilisation of this loop in the PEP-bound *Nme*DAH7PS^{WT} structure, via a salt bridge across the tight dimer interface with Lys107. Inspection of the variation seen for the $\beta_2\alpha_2$ loop indicates that in the apo *Nme*DAH7PS^{Glu176Ala} structure it

adopts a conformation that is different to both the Phe-bound and PEP-bound *Nme*DAH7PS^{WT} (Figure 5.3.4.9).

Although Lys107 adopts a new position in the apo *Nme*DAH7PS^{Glu176Ala} structure, and the partial disorder of its side-chain as well as other residues on the loop is indicative of greater flexibility, a hydrogen bond between Glu98 and Gln172' is evident, like that seen in the Phe-bound *Nme*DAH7PS^{WT} structure. This is interesting, as it seems that in the *Nme*DAH7PS^{Glu176Ala} variant, the absence of the Lys107-Glu176' salt bridge may prime the $\beta_2\alpha_2$ loop for entry into its Phe-bound conformation.

As the loss of interaction between Glu98 and residues of the $\beta_3\alpha_3$ loop and α_3 helix was also observed for the PEP-bound *Nme*DAH7PS^{Gln172Ala}, it may explain the hyper-sensitivity to Phe at low concentrations for both of these variants during kinetic studies as one of the important interactions for the PEP-bound state of *Nme*DAH7PS^{WT} is no longer present. However, this conclusion is speculative and cannot be validated as this structure was obtained in the absence of PEP.

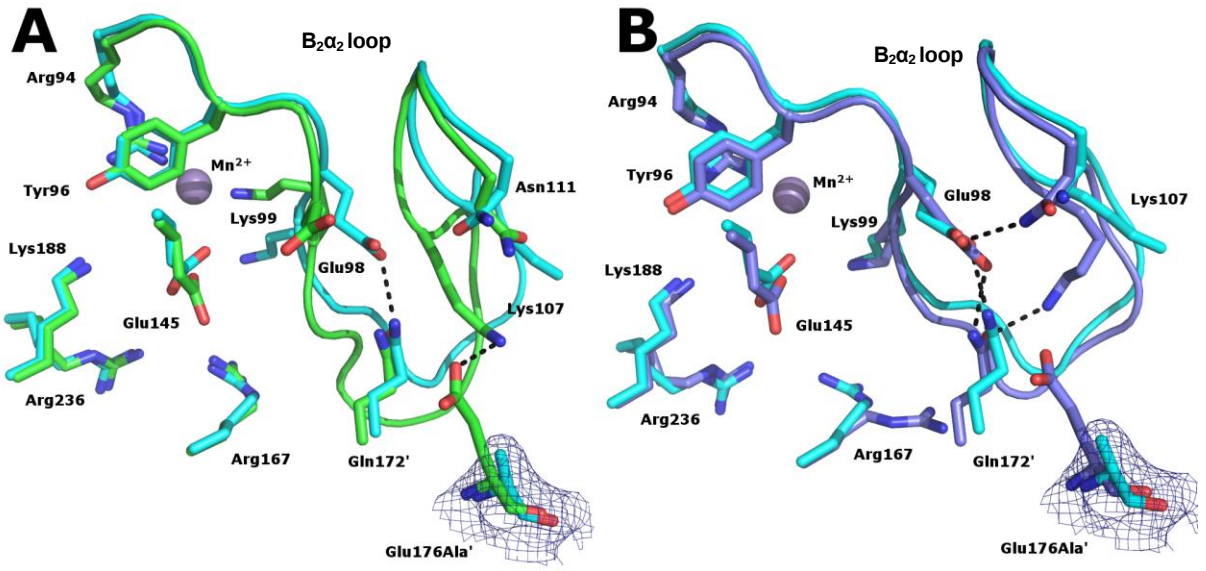


Figure 5.3.4.9 Comparison of the positioning of the $\beta_2\alpha_2$ loop in the apo NmeDAH7PS^{Glu176Ala} (cyan) with NmeDAH7PS^{WT}, illustrating the change in interactions between residues of the $\beta_2\alpha_2$ loop and Gln172 in the NmeDAH7PS^{Glu176Ala} variant. $\lvert 2F_o\text{-}F_c \rvert$ map coloured in blue and contoured at 1.0 σ to illustrate the mutation. A) Comparison between the apo NmeDAH7PS^{Glu176Ala} (cyan) and the PEP-bound NmeDAH7PS^{WT} (green) (PDB code: 4HSN). B) Comparison between the apo NmeDAH7PS^{Glu176Ala} (cyan) and the Phe-bound NmeDAH7PS^{WT} (blue) (PDB code: 4UC5).

The Phe-bound structure was refined to a resolution of 2.04 Å and crystallised in a monoclinic space group ($P12_11$). Data collection and refinement statistics are shown in table 5.3.4.2. The RMSD for the alignment of the tight dimers of the Phe-bound *Nme*DAH7PS^{Glu176Ala} variant to the Phe-bound *Nme*DAH7PS^{WT} structure (PDB code: 4UC5) was 0.347-0.669 Å. This indicates that there is some variation between these two structures, which is localised to the $\beta_2\alpha_2$ loop.

The Phe binding site shows a high degree of structural conservation in the *Nme*DAH7PS^{Glu176Ala} variant (Figure 5.3.4.10A). The $\beta_2\alpha_2$ loop adopts a different conformation to that seen in the Phe-bound *Nme*DAH7PS^{WT} structure in all four chains of the *Nme*DAH7PS^{Glu176Ala} enzyme. In the Phe-bound *Nme*DAH7PS^{WT} structure Gln172' forms a hydrogen bond with both Lys107 and Glu176' indicating that Glu176' may be important for the formation of the Phe-bound hydrogen bonding network, as a hydrogen bond between Lys107 and Gln172' is not formed in the *Nme*DAH7PS^{Glu176Ala} variant enzyme (Figure 5.3.4.10B). The loss of this structural stability, may explain the partial disorder observed for residues of the $\beta_2\alpha_2$ loop in the presence of Phe in *Nme*DAH7PS^{Glu176Ala}.

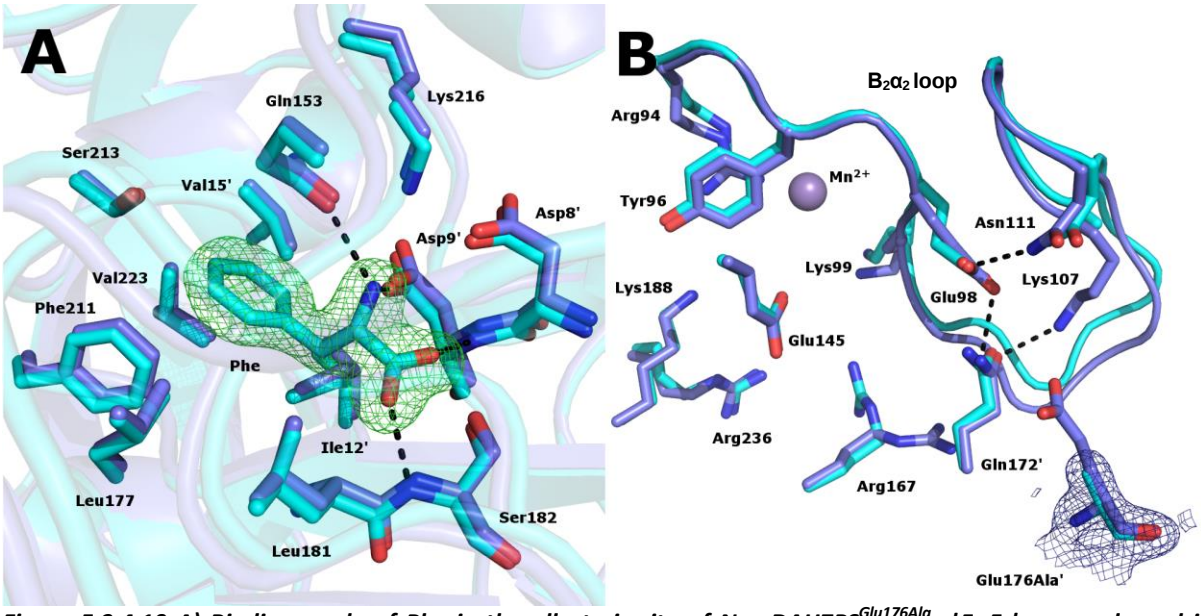


Figure 5.3.4.10 A) Binding mode of Phe in the allosteric site of NmeDAH7PS^{Glu176Ala}. $|F_o-F_c|$ map coloured in green and contoured at 3.0 σ to illustrate presence of Phe in the site at full occupancy. B) Comparison of the positioning of the $\beta_2\alpha_2$ loop in Phe-bound NmeDAH7PS^{Glu176Ala} (cyan) with that of the Phe-bound NmeDAH7PS^{WT} (blue) (PDB code: 4UC5), $|2F_o-F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate the mutation.

The alignment of all four chains of the tetramer illustrates the significant variability in the position of the $\beta_2\alpha_2$ loop. This is indicated in Figure 5.3.4.11, where the residues that make up the $\beta_2\alpha_2$ loop adopt different positions in all four chains and in chain B, four residues of this loop could not be resolved. Considering the resolution (2.04 Å) of this structure and the high degree of disorder in this loop resulting in the limited backbone and side-chain resolvability, there is a high degree of flexibility in this region. The increased flexibility of the $\beta_2\alpha_2$ loop may be correlated to the loss of a hydrogen bond between Lys107 and Gln172′ in the presence of Phe. Noticeably in chain B, for which several residues are not resolved, Arg101 forms a salt bridge with Glu98, like that seen in the *Nme*DAH7PS^{Lys107Ala} variant. This is intriguing, and could possibly aid in the correlation between ITC results between *Nme*DAH7PS^{Lys107Ala} and *Nme*DAH7PS^{Glu176Ala}. The findings for *Nme*DAH7PS^{Glu176Ala} also reiterates what was found for the *Nme*DAH7PS^{Lys107Ala} variant, where the removal of the Lys107-Glu176′ salt bridge across the tight dimer interface by their respective mutations has interesting implications for the conformation adopted by the $\beta_2\alpha_2$ loop, that may be correlated to the ability to form the salt bridge between Arg101 and Glu98. The similarities observed between *Nme*DAH7PS^{Lys107Ala} and *Nme*DAH7PS^{Glu176Ala} variants, further implicates the Lys107-Glu176′ salt bridge and the formation of the Arg101-Glu98 salt bridge in the one-site, largely enthalpically driven binding isotherm obtained for these two variants.

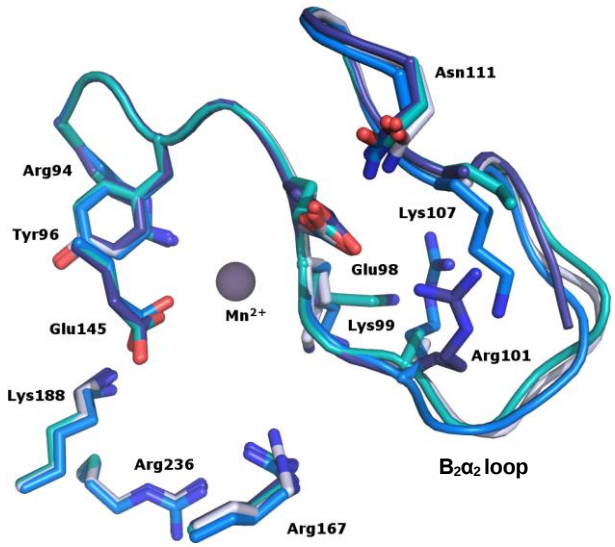


Figure 5.3.4.11 Alignment of all four monomers of the NmeDAH7PS^{Glu176Ala} variant enzyme coloured by chain. Chain A (cyan), Chain B (dark blue), Chain C (sky blue) and Chain D (grey).

Crystal structures of Phe- and PEP-bound NmeDAH7PS^{Glu176Gln}

Two crystal structures of the *Nme*DAH7PS^{Glu176Gln} variant were determined by X-ray diffraction methods. The PEP-bound structure was refined to a resolution of 1.87 Å and crystallised in the same space group as the PEP-bound *Nme*DAH7PS^{WT} ($P12_11$) (PDB code: 4HSN). Data collection and refinement statistics are shown in table 5.3.4.1. The RMSD for the alignment of the tight dimers of the PEP-bound *Nme*DAH7PS^{Glu176Gln} variant to the PEP-bound *Nme*DAH7PS^{WT} structure (PDB code: 4HSN) was 0.151-0.258 Å. This indicates the high degree of similarity between structures. The formation of hydrogen bonds between Gln176' and Lys107, the kind seen in the PEP-bound *Nme*DAH7PS^{WT} are also evident in this crystal structure. This indicates that the same hydrogen bonds can be formed in the PEP-bound structure of this variant upon the change from Glu to Gln (Figure 5.3.4.12). Furthermore, the positioning and conformation of the $\beta_2\alpha_2$ loop remains unchanged in *Nme*DAH7PS^{Glu176Gln}. Ultimately, this indicates that there are no observable structural differences between the PEP-bound *Nme*DAH7PS^{Glu176Gln} and the PEP-bound *Nme*DAH7PS^{WT}, which is in accordance with what was hypothesised for this variant. This observation may also explain the retention of the sequential two-site binding model by ITC.

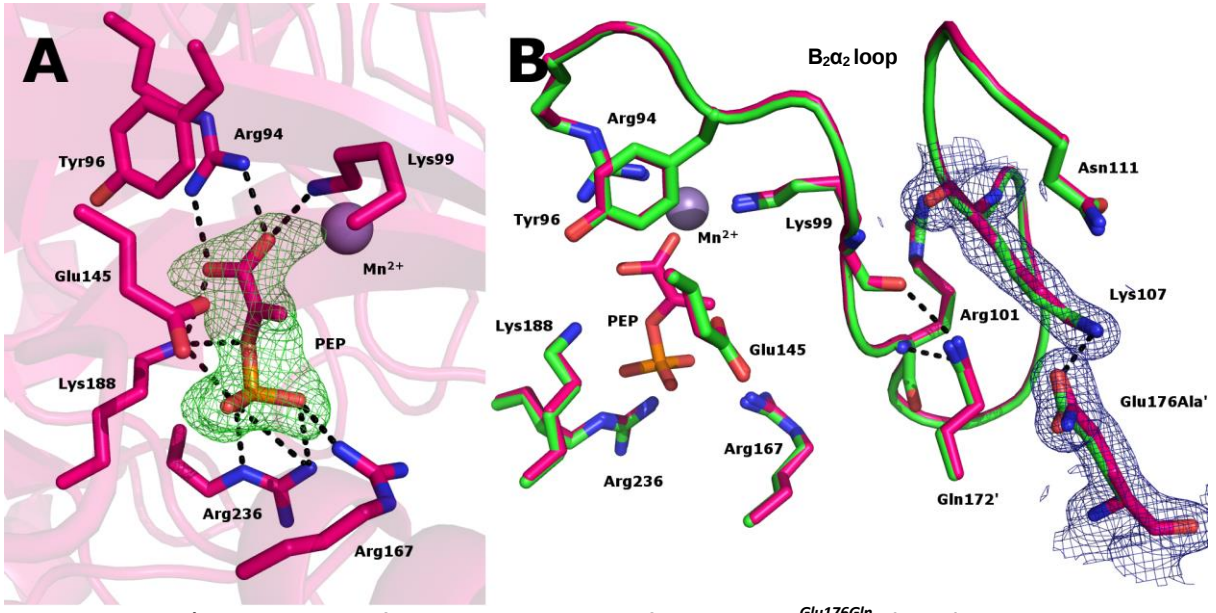


Figure 5.3.4.12 A) Binding mode of PEP in the active site of NmeDAH7PS^{Glu176Gln}. $|F_o\text{-}F_c|$ map coloured in green and contoured at 3.0 σ to illustrate presence of PEP in the site at full occupancy. B) Similarity in conformation of the $\beta_2\alpha_2$ loop in NmeDAH7PS^{Glu176Gln} (pink) with that of the NmeDAH7PS^{WT} (green) (PDB code: 4HSN) illustrating the retention of hydrogen bonds between residues of the $\beta_2\alpha_2$ loop and residues: Gln172 and Gln176 in the NmeDAH7PS^{Glu176Gln} variant. $|2F_o\text{-}F_c|$ map coloured in blue and contoured at 1.0 σ to illustrate the hydrogen bond between Lys107 and Gln176.

The Phe-bound crystal structure of the *Nme*DAH7PS^{Glu176Gln} was refined to a resolution of 2.70 Å and crystallised in the same space group as the PEP-bound *Nme*DAH7PS^{WT} structure (PDB code: 4HSN) ($P12_11$). Data collection and refinement statistics are shown in table 5.3.4.2. The RMSD for the alignment of the tight dimers of the Phe-bound *Nme*DAH7PS^{Glu176Gln} variant to the Phe-bound *Nme*DAH7PS^{WT} structure (PDB code: 4UC5) was 0.218-0.287 Å. This indicates large structural similarity between these structures. However, the poor resolution of this structure limits the confidence in identification of structural variation in this variant.

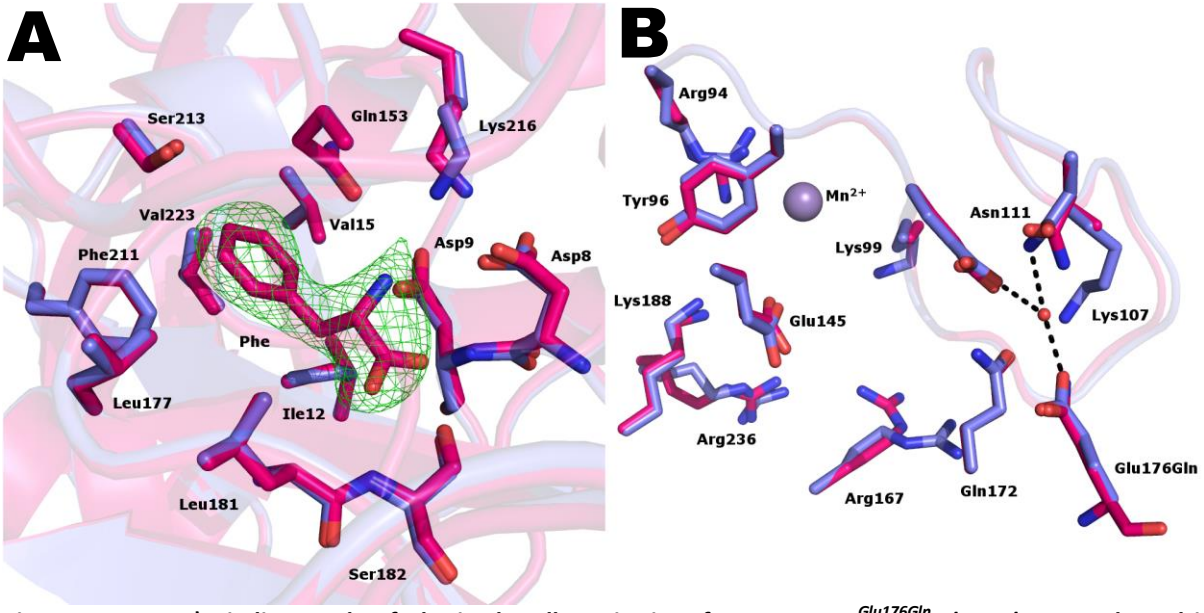


Figure 5.3.4.13 A) Binding mode of Phe in the allosteric site of NmeDAH7PS^{Glu176Gln}. $|F_o-F_c|$ map coloured in green and contoured at 3.0 σ to illustrate presence of Phe in the site at full occupancy. B) Comparison of the positioning of the $\beta_2\alpha_2$ loop in Phe-bound NmeDAH7PS^{Glu176Ala} (pink) with that of the Phe-bound NmeDAH7PS^{WT} (blue) (PDB code: 4UC5) illustrating the new conformation of Asn111 and its loss of hydrogen bond with Glu98. The conserved water that links residues 98, 111 and 176 in the NmeDAH7PS^{WT} structure is shown as a small red sphere.

Inspection of the Phe binding site of the *Nme*DAH7PS^{Glu176Gln} variant indicates the structural similarity to the allosteric site of the Phe-bound *Nme*DAH7PS^{WT} structure. The residues surrounding the $\beta_2\alpha_2$ loop illustrate the lack of structural change in the *Nme*DAH7PS^{Glu176Gln} variant. Interestingly, Asn111 in this structure does form a hydrogen bond with Glu98 in this variant enzyme (Figure 5.3.4.13). The conserved water molecule that links Glu98, Asn111 and Glu176 in two chains of the Phe-bound and in all four chains of the double liganded *Nme*DAH7PS^{WT} structures (PDB codes: 4UC5 & 5DCB), could not be resolved in the *Nme*DAH7PS^{Glu176Gln} structure and may again be a result of the low resolution of this structure. Furthermore, inferences based on subtle variations in the positions of side-chains are also

dubious at this resolution. Overall, this variant enzyme shows little difference to *Nme*DAH7PS^{WT}, which is perhaps not surprising due to the subtle change in the side-chain of Glu176′.

	***Nme*DAH7PS Glu98Ala**	***Nme*DAH7PS Glu98Ala**	***Nme*DAH7PS Lys107Ala**	***Nme*DAH7PS Gln172Ala**	***Nme*DAH7PS Gln172Ala**
	L-Phenylalanine	**PEP**	**L-Phenylalanine and PEP**	**L-Phenylalanine**	**PEP**
Data Collection					
Crystal system; space group	Monoclinic, $P12_11$	Monoclinic, $P12_11$	Monoclinic, $P12_11$	Monoclinic, $P12_11$	Monoclinic, $P12_11$
Unit cell parameters					
a, b, c (Å)	73.32, 142.33, 74.59	73.85, 135.25, 76.04	75.33, 141.03, 76.03	73.17, 141.85, 74.76	73.07, 138.02, 76.27
α, β, γ (º)	90, 96.84, 90	90, 95.95, 90	90, 97.02, 90	90, 96.40, 90	90, 96.03, 90
Resolution range (Å)	74.06-2.42 (2.47-2.42)	75.63-2.50 (2.56-2.50)	74.46-1.75 (1.78-1.75)	47.28-2.30 (2.35-2.30)	75.85-2.68 (2.78-2.68)
Measurements	440483	335301	1195557	505973	158867
Unique reflections	57783	50068	158105	67189	42187
Redundancy	7.6 (7.1)	6.7 (6.8)	7.6 (7.4)	7.5 (7.6)	3.8 (3.8)
Completeness (%)	99.8 (97.1)	97.7 (97.6)	100.0 (99.7)	100.0 (100.0)	99.8 (100)
I/σ (I)	11.0 (1.76)	9.0 (2.27)	16.9 (2.14)	14.6 (2.3)	12.8 (1.7)
R_{merge}	0.156 (0.906)	0.122(0.753)	0.079 (0.869)	0.094 (0.870)	0.069 (0.619)
$CC_{1/2}$	0.702	0.784	0.68	0.716	0.707
Wilson B value $(Å)^2$	31.4	40.77	20.70	36.65	53.96
Matthews coefficient	2.54 (51.65)	2.48 (50.53)	2.64 (53.38)	2.54 (51.54)	2.52 (51.15)
Refinement					
R_{work}	0.1910	0.2003	0.1495	0.1859	0.1956
R_{free}	0.2457	0.2639	0.1756	0.2319	0.2586
Chain length	351	351	351	351	351
Observed number of residues	345 (Chain A, B & C) 346 (Chain D)	333 (chain A) 332 (Chain B, C &D)	345 (Chain A, B & C) 346 (Chain D)	345 (Chain A, C &D) 343 (Chain B)	332 (Chain A, B & C) 333 (Chain D)
Water molecules	331	112	1132	232	14
Other	8 (Mn^{2+}, EDO, PEG)	8 (Mn^{2+}, SO_4^{2-}, EDO)	15 (Cl^-, SO_4^{2-}, PEG, Mn^{2+})	10 (EDO, PEG, SO_4^{2-}, Mn^{2+})	4 (Mn^{2+}, EDO)
Ligand	8 (PEP(0.75 occ.), Phe)	4 (PEP)	8 (PEP, Phe)	4 (Phe)	4 (PEP)
Mean B $(Å)^2$					
Protein	35.55	47.79	23.59	47.91	63.65
Water	30.13	37.02	29.96	42.47	42.10
Other	43.87	74.15	57.68	71.73	66.03
Ligand	46.66, 26.38	64.35	21.65, 17.71 (PEP, Phe)	39.17	87.17
R.m.s.d from target values					
Bond lengths (Å)	0.0101	0.0111	0.0103	0.0097	0.0088
Bond angles	1.3819	1.4309	1.3868	1.3300	1.2587
Dihedral angles	0.0721	0.0743	0.0826	0.0738	0.0679
Ramachandran					
Preferred (%)	96.48	95.54	97.62	96.87	95.73
Allowed (%)	3.37	4.16	2.16	2.98	4.04
Outliers (%)	0.15	0.30	0.22	0.15	0.23
PDB Entry	5CZS	5CZ0	5D05	XXXX	XXXX

Table 5.3.4.1 Data collection and refinement statistics for NmeDAH7PS^{Glu98Ala} and NmeDAH7PS^{Gln172Ala} in the presence and absence of Phe as well as NmeDAh7PS^{Lys107Ala} in the presence of both PEP and Phe. (PDB codes: 5CZS, 5CZ0, 5D05, XXXX and XXXX) (XXXX-un-deposited structure) .

	***Nme*DAH7PS Glu176Ala**	***Nme*DAH7PS Glu176Ala**	***Nme*DAH7PS Glu176Gln**	***Nme*DAH7PS Glu176Gln**
	L-Phenylalanine	**Apo**	**L-Phenylalanine**	**PEP**
Data Collection				
Crystal system; space group	Monoclinic, $P12_11$	Orthorhombic, $P2_12_12_1$	Monoclinic, $P12_11$	Monoclinic, $P12_11$
Unit cell parameters				
a, b, c (Å)	75.65, 140.83, 77.27	78.38, 135.37, 74.68	73.43, 139.70, 75.12	73.54, 135.55, 76.01
α, β, γ (º)	90, 98.56, 90	90, 90, 90	90, 95.83, 90	90, 96.27, 90
Resolution range (Å)	76.41-2.04 (2.08-2.04)	100.47-2.40 (2.46-2.40)	46.89-2.70 (2.81-2.70)	75.55-1.87 (1.90-1.87)
Measurements	772315	922106	312591	455196
Unique reflections	101442	62994	41374	121730
Redundancy	7.6 (7.2)	14.6 (14.0)	7.6 (7.7)	3.7 (3.8)
Completeness (%)	99.9 (98.7)	100.0 (99.9)	99.9 (99.9)	99.9 (100.0)
I/σ (I)	18.6 (2.6)	19.5 (2.3)	17.4 (2.2)	11.7 (1.77)
R_{merge}	0.079 (0.696)	0.122 (1.21)	0.074 (1.014)	0.054 (0.890)
$CC_{1/2}$	0.808	0.772	0.851	0.600
Wilson *B* value $(Å)^2$	28.40	41.08	64.12	31.20
Matthews coefficient	2.68 (54.09)	2.62 (53.01)	2.52 (51.25)	2.48 (50.38)
Refinement				
R_{work}	0.1706	0.2006	0.2182	0.1761
R_{free}	0.2057	0.2557	0.26835	0.2207
Chain length	351	351	351	351
Observed number of residues	345 (All four chains)	332 (Chains A, C &D), 333 (Chain B)	343 (Chains A, B & C), 344 (Chain D)	333 (Chain A, C &D) 334 (Chain B)
Water molecules	745	230	17	640
Other	7 (Cl^-, Mn^{2+})	8 (EDO, Mn^{2+}, SO_4^-)	7 (EDO, Mn^{2+}, Cl^-)	11 (Mn^{2+}, EDO, SO_4^{2-})
Ligand	4 (Phe)	-	4 (Phe)	4 (PEP)
Mean *B* $(Å)^2$				
Protein	31.58	42.13	85.44	38.53
Water	34.50	34.95	81.91	40.80
Other	37.79	55.65	84.77	57.57
Ligand	22.77	N/A	68.93	38.47
R.m.s.d from target values				
Bond lengths (Å)	0.0107	0.0118	0.0095	0.0114
Bond angles	1.3910	1.4816	1.2953	1.4502
Dihedral angles	0.0803	0.0799	0.0668	0.0863
Ramachandran				
Preferred (%)	97.06	96.68	96.71	97.15
Allowed (%)	2.35	3.02	3.00	2.63
Outliers (%)	0.59	0.30	0.29	0.22
PDB Entry	5CZT	XXXX	XXXX	5D02

Table 5.3.4.2 Data collection and refinement statistics for NmeDAH7PS^{Glu176Ala} and NmeDAH7PS^{Glu176Gln} in the presence and absence of Phe. (PDB codes:5CZT, XXXX, XXXX and 5D02) (XXXX-un-deposited structure).

5.4 DISCUSSION

Several variants of *Nme*DAH7PS were generated to better our understanding of how *Nme*DAH7PS communicates allostery between its active and allosteric sites. Two mutants were designed to elucidate the importance of hydrophobic interactions within the allosteric site. A further five variant enzymes were designed to understand the importance of hydrogen bonding network reorganisation in the transfer of allosteric information.

Hydrophobic binding pocket has important role in Phe binding

The structural and functional characterisation of the two allosteric site variant enzymes has reiterated the importance of hydrophobic interactions in the allosteric binding pocket for ligand recognition and binding. The findings illustrate the importance of the β-hairpin insertion for allostery, on which both Phe211 and Val223 are located. Both of these residues were shown by kinetics to be crucial to the sensitivity of *Nme*DAH7PS to Phe and this reduction or loss of sensitivity to Phe appears to be due primarily to a reduction or loss of binding affinity for Phe.

Contrary to previous studies, this investigation illustrated the importance of Phe211 for the binding of Phe. It has been shown that the edge to face π-stacking interaction provided by Phe211, or the structural organisation of the allosteric site provided by the aromatic ring of Phe211 prior to Phe binding are crucial to Phe binding in *Nme*DAH7PS.

Crucial connection to allosteric site within same chain is important for regulation

The active site of *Nme*DAH7PS is connected to the allosteric site within the same chain, via Glu98 located at the N-terminal end of the $\beta_2\alpha_2$ loop. Glu98 interacts with residues: Asp148, Thr151 and Tyr155 from the $\beta_3\alpha_3$ loop and α_3 helix. Gln153, also located on $\beta_3\alpha_3$ loop has a role in binding the amine moiety of Phe. Mutation of Glu98 resulted in a decreased K_M for the substrate PEP and implies that the interaction, which connects allosteric and active sites within

the same chain may compromise the apparent affinity of *Nme*DAH7PS^{WT} for its natural substrate. In the *Nme*DAH7PS^{Glu98Ala} variant, PEP was bound in its reactive conformation in the presence of Phe, unlike that observed for *Nme*DAH7PS^{WT}. This suggests that Glu98 may be involved in the inversion of the binding mode of PEP, and is an important contributor to the allosteric mechanism of *Nme*DAH7PS^{WT}. This finding is probably due to the loss of the ability for this residue to take part in the postulated hydrogen bonding network via interactions with Lys107, Asn111 and Gln172′. The involvement of this residue in the allosteric mechanism of *Nme*DAH7PS is proven by the significant decrease in sensitivity to Phe at all concentrations tested for this variant.

Interactions across the tight dimer interface are crucial for the allosteric mechanism o *Nme*DAH7PS

There is a notable interaction across the tight dimer interface in the PEP-bound enzyme consisting of a salt bridge between Lys107 of the $\beta_2\alpha_2$ loop and Glu176′ of the α_4 helix. Both of these residues were mutated in this study, with interesting results. Unfortunately, the *Nme*DAH7PS^{Lys107Ala} variant was catalytically inactive, due to the retention of Phe, revealed by ITC and crystallographic studies. The binding isotherm was however fit to a one-site binding model, a feature that was also evident for *Nme*DAH7PS^{Glu176Ala}. The thermodynamic parameters indicated a very large enthalpic and minimal entropic contribution to binding in both of these variants.

Taken together, the binding isotherms for these two variants suggest that the Lys107-Glu176′ salt bridge is a contributor to the entropic penalty paid upon Phe binding to *Nme*DAH7PS^{WT}. Therefore in the absence of the Lys107-Glu176′ salt bridge, in these variants, the entropic penalty has already been paid, thus reducing the entropic contribution to the binding isotherm. Crystal structures of these variants highlighted an important feature that may result from the loss of the Lys107-Glu176′ salt bridge. In the presence of Phe, a salt bridge between Arg101 and Glu98 was evident in all four chains of *Nme*DAH7PS^{Lys107Ala} and in one chain of *Nme*DAH7PS^{Glu176Ala}. This new interaction, within the $\beta_2\alpha_2$ loop may contribute to the enthalpic character of the binding isotherm in these variants.

The high degree of disorder in the $\beta_2\alpha_2$ loop when Phe is present at full occupancy in *Nme*DAH7PS^{Glu176Ala} may correlate to the limited inhibition achieved by Phe at concentrations > 200 µM. The hyper-sensitivity to Phe at low concentrations, observed for the *Nme*DAH7PS^{Glu176Ala} variant suggests that the loss of stabilising interactions between Glu176′ and the $\beta_2\alpha_2$ loop in the absence of Phe may prime the variant for inhibition by Phe.

The correlation between these two variants experimentally illustrates the importance of the Lys107-Glu176′ salt bridge in propagation and mediation of the allosteric signal, as well as the formation of the active and Phe-bound interaction networks around the $\beta_2\alpha_2$ loop in *Nme*DAH7PS.

It was postulated that Gln172′, located on the same α_4 helix as Glu176′, would have a crucial involvement in the allosteric mechanism of *Nme*DAH7PS^{WT}. Gln172′ forms interactions with Lys99 and Arg101 of the conserved KPR(T/S) motif in the PEP-bound *Nme*DAH7PS^{WT} structure and both Glu98 and Lys107 of the $\beta_2\alpha_2$ loop in the Phe-bound *Nme*DAH7PS^{WT} structure. However, the role of Gln172′ in the reorganisation of the $\beta_2\alpha_2$ loop in the presence of Phe was not found to be the case and is in agreement with MD simulations, which do not highlight it as an important residue in the formation of the Phe-bound network. The lack of involvement of Gln172′ during regulation by Phe was exposed by the Phe-bound crystal structure of the *Nme*DAH7PS^{Gln172Ala} variant, where the $\beta_2\alpha_2$ loop adopted the same conformation as that observed in the Phe-bound *Nme*DAH7PS^{WT} structure. Its involvement with the $\beta_2\alpha_2$ loop during catalysis however, was found to be far more revealing. The catalytic rate and K_M for both substrates increased in the *Nme*DAH7PS^{Gln172Ala} variant. The PEP bound crystal structure of this variant indicated that partial disorder for the $\beta_2\alpha_2$ loop was evident three of four chains. This observation suggests that Gln172′ may have a key role in determining the conformations available to the $\beta_2\alpha_2$ loop.

The hyper-sensitivity to low concentrations of Phe, observed for both the *Nme*DAH7PS^{Gln172Ala} and *Nme*DAH7PS^{Glu176Ala} may be a result of the loss of interaction with the $\beta_2\alpha_2$ loop observed in their respective Phe-less crystal structures. The loss of these interactions with the $\beta_2\alpha_2$ loop

in the absence of Phe suggests that these variants may be primed for inhibition by Phe. The differences in catalytic rate observed for these variants implies that this feature unlikely to be a direct result of absolute activity.

Conclusion

The complicated network of interactions between residues of the $\beta_2\alpha_2$ loop across the tight dimer interface with the β-hairpin insertion and α_4 helix and within the same chain with residues of the $\beta_3\alpha_3$ loop and α_3 helix have been identified in this study. The experimental results have a high level of redundancy due to the complexity of these interaction networks. However, the extensive description of the structural and functional relationships between the variants involved in this study has illustrated the importance of these residues to the propagation and conferral of allosteric signal. It has been shown that careful control of the dynamics of the $\beta_2\alpha_2$ loop by residues on or surrounding it are crucial to both catalysis and regulation in *Nme*DAH7PS.

CHAPTER 6- SUMMARY AND FINAL REMARKS

This thesis describes a detailed investigation of the structure and function of *Nme*DAH7PS. This has involved a detailed study of the active site of *Nme*DAH7PS as a as a potential target for the development of novel antibiotics against bacterial meningitis. Investigations into the quaternary structure of *Nme*DAH7PS have also identified the importance of the tetrameric assembly of this enzyme for catalysis and allostery. Perhaps most importantly, this thesis explains in detail the subtleties of the allosteric mechanism adopted by this enzyme in the absence of large conformational change. This study has revealed that allostery is mediated by the interactions between the $\beta_2\alpha_2$ loop and surrounding ionisable residues of the tight dimer.

This chapter entails a summary of the key findings from the previous chapters and identifies the avenues for future work on this enzyme. It also illustrates the applicability of this work to other systems in nature.

6.1 Stereochemistry of inhibitor enables accommodation of nucleophilic water, which is crucial for potency

The structural and functional analysis of three highly ligand-efficient PEP-mimicking inhibitors of *Nme*DAH7PS has revealed the important characteristics required for potency. These characteristics have been shown to be in accordance with the stereochemistry of the catalytic mechanism and are centred on the nucleophilic attack of water at C_2 of the oxocarbenium ion during catalysis. The conformation of the facial selectivity of the nucleophilic attack has been identified by the correlation between inhibitor potency and the favourable hydrogen bonding geometry and optimal van der Waals interactions with an active site water molecule in close proximity to C_2 of all inhibitor and substrate bound structures (PDB codes: 4HSN, 4UMA, 4UMB and 4UMC).[26] However, the active site of *Nme*DAH7PS has been shown to most effectively accommodate a planar mimic or intermediate (*E*-vinyl phosphonate), implying that the planar oxocarbenium ion is an important high energy intermediate in catalysis.[68–70] Based on transition state theory, an enzyme will bind most strongly to the transition state, therefore an enzyme that is able to mimic characteristics of this state will have the greatest potency.[69] Accommodation of this active site water molecule by this inhibitor and a stronger coulombic interaction with Lys99 are believed to be the main contributors to the inhibitor potency.

This is likely to provide a platform for future work in this area, which would involve the extension of this highly atom-efficient inhibitor[137] into a seven carbon oxocarbenium intermediate mimic containing this planar phosphonate functionality at C_2 and a rotatable isosteric phosphonate or phosphate functionality for occupation of the E4P binding site, similar to the bisphosphate inhibitors.[54,138] This work will also involve the addition of backbone hydroxyl moieties in order to obtain more contacts within the site as they have been shown to be important for the affinity of E4P within the active site of *Eco*DAH7PS.[49,53] This synthetic scheme would involve a carbohydrate scaffold of correct stereochemistry as the starting point (Figure 6.1.1).

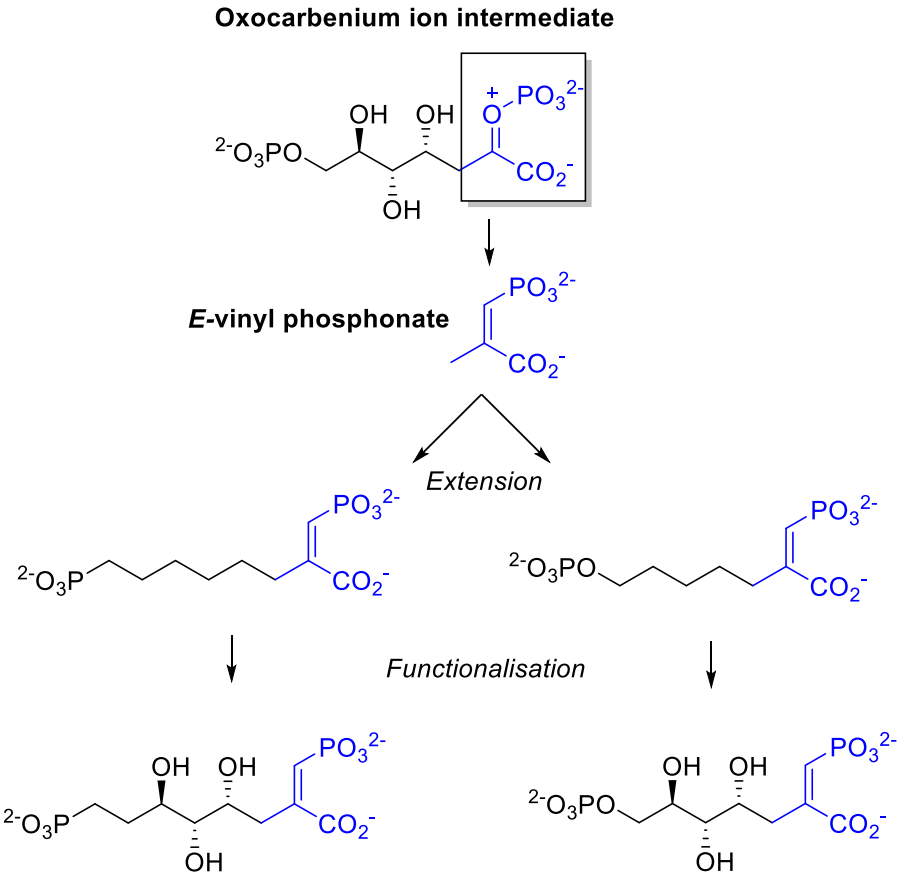


Figure 6.1.1 Extension and functionalisation of the highly atom efficient E-vinyl phosphonate inhibitor, oxocarbenium ion intermediate mimic

Furthermore, given this structural understanding of the PEP binding site and the inhibitory potency dictated by interactions at this site, it is possible to computationally screen the site with a library of small ligands in order to identify new optimal contacts within the *Nme*DAH7PS active site. The extensive work previously done in the area of fragment-based-ligand-discovery

(FBLD) involving library screening, compound selection and experimental testing, including: STD-NMR, Water LOGSY, DSF, ITC and crystallography will provide great resources in these attempts.[139–147] The extensive knowledge of the *Nme*DAH7PS active site as well as the stability and proven crystallisation conditions would bode well in the deduction of structure-activity relationships (SAR). The culmination of this approach with that of the reaction intermediate based drug discovery previously mentioned may lead to the design and synthesis of an inhibitor that appropriately accommodates a large percentage of the chemical space present within the site.

6.2 The allosteric site of *Nme*DAH7PS is selective for aromatic amino acids, in particular, Phe

The aromatic amino acid, Phe was found to be the most potent feedback inhibitor of *Nme*DAH7PS and this potency was highly correlated to binding affinity. The important interactions distinguishing binding affinity between these aromatic amino acids are localised to the β_{6a} strand. The steric hindrance generated by the hydroxyl of Ser213, was found to act as a selectivity filter in this enzyme, negating the binding of Tyr, which is in agreement with previous work.[26] The recruitment of Phe211 and its π-stacking interaction with the aromatic ring of the amino acid were identified as important contributors to the binding affinity of the aromatic amino acid.

Nonetheless, it is apparent that this site is extremely selective for the binding of these aromatic amino acids, in that, the movement of the β-hairpin insertion is linked to the hydrophobicity of the aromatic ligand. Furthermore, the N-terminal “allosteric binding loop” is highly selective for the amino carboxylate functionality, forming several electrosatic interactions. The allosteric site exhibits a high selectivity for these aromatic amino acids and lack of promiscuity between them. Therefore, attempts to target this site with effective alternative inhibitors would prove difficult.

While *Neisseria meningitidis* expresses a single DAH7PS, other organisms express multiple paralogues that are each sensitive to feedback inhibition by a single aromatic amino acid. *E. coli* and *S. cerevisiae* both exhibit this characteristic; however, *S. cerevisiae* lacks the Trp sensitive paralogue. Interestingly, the Trp sensitive DAH7PS from *E. coli* has not had any structural or functional investigation. Given the understanding of the selectivity between Phe

and Tyr in these organisms, the characteristics that make up a Trp selective site are not well understood. A full investigation into the Trp sensitive paralog from *Eco*DAH7PS could give valuable insight into the sensitivity of *Nme*DAH7PS to high concentrations of Trp.

6.3 Hydrophobic interactions are important for allostery in *Nme*DAH7PS

Two variant *Nme*DAH7PS enzymes were made to better understand the ligand recognition of the largely hydrophobic allosteric site. Both of these point mutations (*Nme*DAH7PS^{Phe211Ala} and *Nme*DAH7PS^{Val223Ala}) are located on the β-hairpin insertion and were shown to be important for ligand binding. Phe211 was shown to be crucial for Phe to bind at any concentration, whereas substitution of Val223 exhibited reduced affinity for Phe.

This work has focused on residues important for ligand recognition; however, future work should aim to identify the contribution of hydrophobic interactions to the communication of allosteric information in *Nme*DAH7PS. The work in this thesis on the transduction of an allosteric signal has focused entirely on ionisable and hydrogen bonding residues, but, the contribution of these hydrophobic residues to the transferal of allosteric signal should not be discounted. There are several hydrophobic residues that may be important for the conferral of regulation in *Nme*DAH7PS including but not limited to: Phe119 on the $\beta_2\alpha_2$ loop, Phe144 on the $\beta_3\alpha_3$ loop, Tyr155 on the α_3 helix and Leu177 and Leu181 on the α_4 helix. Successive mutations of these residues and their characterisation may illustrate their role in the allosteric response of *Nme*DAH7PS.

6.4 Allosteric inhibitor binding pre-organises active site to bind PEP in a flipped (un-reactive) orientation.

Crystallographic studies have shown that in the presence of allosteric effector, the active site of *Nme*DAH7PS undergoes a structural reorganisation. Such reorganisation, which occurs at high enough concentrations for all ligands (Phe, Trp and Tyr) is associated with a change in the binding mode of PEP, resulting in a “flipped” and unreactive binding orientation. This active site reorganisation is the same in the presence and absence of PEP and suggests that the site is pre-organised for binding PEP in this inverted orientation.

The reorganisation of the site also has implications for the other substrate E4P, where the new conformation adopted by residue Lys99, resulted in the formation of a salt-bridge with the metal coordinating residue, Asp324. This interaction may hinder the coordination of the aldehyde functionality of E4P to the divalent metal ion. Further investigations into the importance of this interaction would involve targeting the residues that may contribute to this flipped orientation in the enzymes regulated state. It has already been shown for the *Nme*DAH7PS^{Glu98Ala} variant, that PEP adopted a normal (active) binding mode in the presence of Phe. Therefore, it can be inferred that the residues of the $\beta_2\alpha_2$ loop and their interactions with surrounding residues are important for the new "flipped" binding mode of PEP.

ITC experiments, involving the binding of the substrate PEP in the presence and absence of Phe may give a thermodynamic understanding of substrate binding to *Nme*DAH7PS in its inhibited state. There is likely to be changes in the binding isotherm, which may give insight into the mechanism of PEP binding in the presence of Phe.

6.5 Treatment of the dimer as the functional unit of *Nme*DAH7PS

The presence of a tetrameric interface in *Nme*DAH7PS was found to be unnecessary for catalytic function and regulation. The removal of a salt bridge (Arg126-Glu27) at the tetramer interface resulted in a dimeric *Nme*DAH7PS that had similar kinetic and Phe binding characteristics to the wild-type enzyme. Furthermore, removal of the only interaction (His219-His219) at the Chain A-D and B-C interface also resulted in an enzyme with similar characteristics to the wild-type enzyme. The similarity between these two variant enzymes and the *Nme*DAH7PS^{WT} enabled treatment of *Nme*DAH7PS as a functional dimer for the experiments described in chapter 5.

An interesting finding from these investigations was the potential role of Glu17 in the closure of the N-terminal tail over the bound ligand. Glu17, located slightly towards the N-terminus from Glu27 (part of the salt bridge) and forms a hydrogen bond with His219 located on the $\beta_{6a}\beta_{6b}$ loop in the presence of Phe. The interaction of Glu17 with His219, implicates it in the closure of the allosteric binding loop and the communication between chains of the tight dimer. Characterisation of this variant may provide further insight into importance of the interactions between chains of tight dimer for the regulation of *Nme*DAH7PS

6.6 Entropically driven method of allostery employed by *Nme*DAH7PS

The isotherm for binding of Phe to *Nme*DAH7PS is indicative of an entropically driven allosteric ligand binding event. The increased flexibility of *Nme*DAH7PS in the presence of Phe has also been observed in the MD simulations, where an increase in the amplitude and a decrease in frequency of atomic motions was evident for large regions of the enzyme ($\beta_2\alpha_2$ loop, $\beta_3\alpha_3$ loop, α_3 helix and β-hairpin insertion) as well as the loss of correlated motions between residues in the Phe-bound state. This increase in flexibility is evident despite the ordered nature of the allosteric binding loop upon Phe binding. The correlation between these two methods has interesting implications for the method of allostery in *Nme*DAH7PS.

Further work in this area would involve trying to understand the flexibility incurred on *Nme*DAH7PS upon Phe binding using NMR relaxation experiments, however, even the dimeric variant remains very large for such a technique at this point. This technique may provide insight into the conformational entropy changes in this enzyme upon Phe binding and may reveal the local contributions to the changes in heat capacity.[148] Another method that could be used to better understand this hypothesis would be to use ITC methods at a variety of temperatures to obtain the heat capacity (ΔC_p) changes in *Nme*DAH7PS during Phe binding.[149,150] A combination of these methods, with the aid of the MD and crystallographic studies would provide valuable insight to the entropically driven allostery adopted by *Nme*DAH7PS.

6.7 Ionisable side-chains mediate the allosteric signal in *Nme*DAH7PS

A method of understanding the mechanism of allostery adopted by *Nme*DAH7PS has involved the structural (X-ray crystallography), computational (MD) and pH dependence studies. These studies have illustrated the importance of ionisable residues in the catalysis and regulation of *Nme*DAH7PS. Initial structural studies described the reorganisation of electrostatic and hydrogen bonding partners in the presence of allosteric ligand. This hypothesis was investigated by MD studies, which also highlighted other residues with potential roles in the formation of the network that were unobservable from the structural analysis. This is due to the fact that MD illustrates protein dynamics, rather than the static picture that is delivered by the crystal structure.

The involvement of these residues in the allosteric network was identified via monitoring the pK_a fluctuations of ionisable residues in the protein, thus giving an understanding of the changes in the local environment of residues upon Phe binding. The allosteric effect of Phe was shown to be dependent on the pH at which the kinetic measurements were carried out. This observation implies that limitation of the pK_a changes of the highlighted residues required for communication of allosteric information in *Nme*DAH7PS can be achieved changing the pH. This finding showed that networks, involving ionisable residues are involved in transmission of the allosteric signal to the active site.

This methodology, while focusing on a small selection of residues, gives insight into the dynamics of *Nme*DAH7PS during allostery. However, the methods described are not in solution (crystal packing effects) and/or are purely computational (possibly biased by the crystal structure conformation), therefore it would be beneficial to get an understanding of the dynamics of *Nme*DAH7PS experimentally. The use of hydrogen/deuterium exchange mass spectrometry experiments may be revealing for this investigation. This techniqe may give an experimentally validated understanding of the interactions between residues in solution and shed light on the dynamics of these interactions predicted by MD.[151]

6.8 Regulatory effects of Phe binding occur both within the same chain and across the tight dimer interface

These investigations have highlighted several residues that are important for the transmission of allosteric signal. The transmission of this signal has been shown to propagate via two general pathways (Figure 7.8). The first, within the same chain, consists of interactions between Glu98 of the $\beta_2\alpha_2$ loop and Asp148, Thr151 and Tyr155 of the $\beta_3\alpha_3$ loop and α_3 helix respectively. The importance of this set of connections has been illustrated in the *Nme*DAH7PS^{Glu98Ala} variant, where the allosteric response to Phe binding was severely attenuated.

The second pathway is far more complex and extends across the tight dimer interface. It involves interaction of Lys107 of the $\beta_2\alpha_2$ loop with Gln172′ and Glu176′, of the α_4 helix. Lys 107 also interacts with residues of the β_{6a} strand (His209′ and the peptide amino group of His210′, adjacent to Phe211′). The removal of the side-chains of some of these residues had a

variety of effects on the catalysis and regulation of *Nme*DAH7PS. The high redundancy encountered when making single point variant enzymes makes delineation of the exact network challenging. However, it has been shown that the Lys107-Glu176′ salt bridge is crucial for modulating the binding affinity of Phe as well as the cooperativity observed in the binding isotherm between allosteric sites of the tight dimer. Crystallographic studies have implicated the formation of a salt bridge between Arg101 and Glu98 in this high affinity for Phe. This feature is observed in the Phe-bound crystal structure of the Phe-sensitive paralogue from *E. coli*.[101]

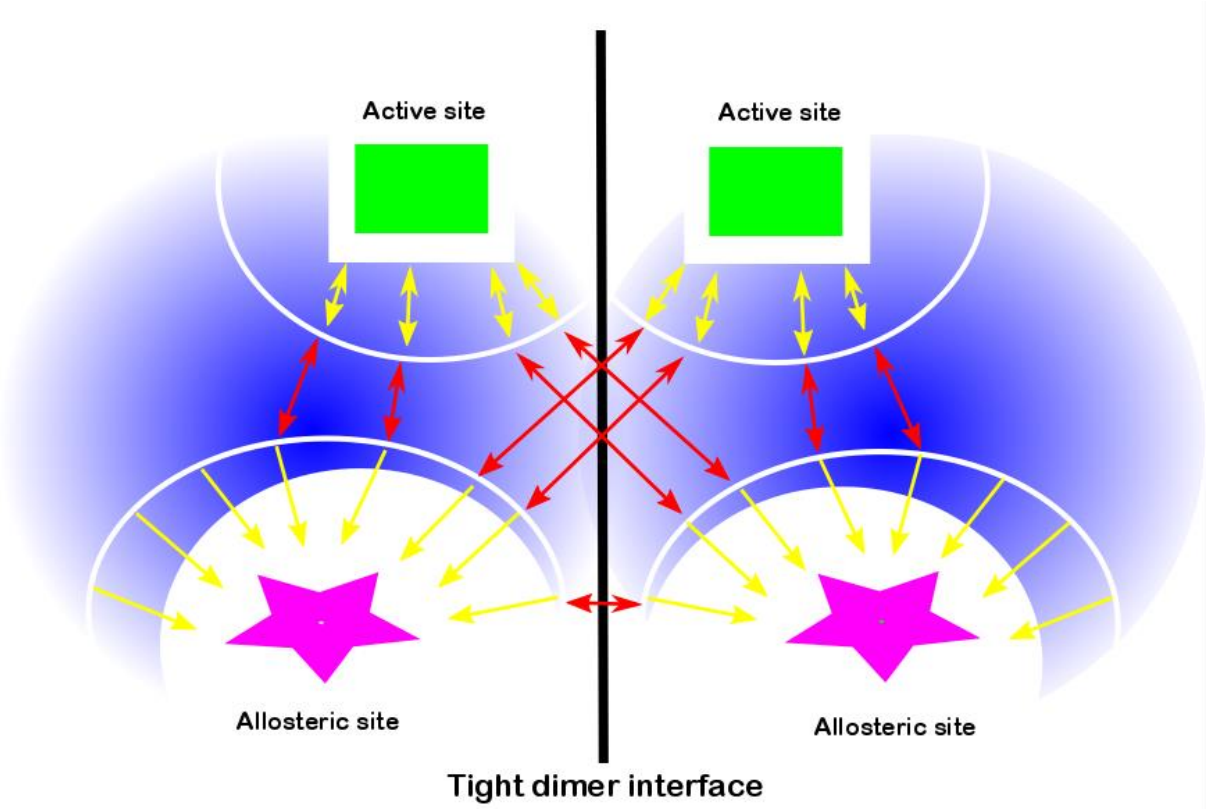


Figure 7.8.1 The communication of allostery both within the same chain and the adjacent chain via the tight dimer interface. Red arrows indicate communication of allostery. Yellow arrows indicate ligand binding. Allosteric Phe shown as magenta star and the active site ligand is shown as a green rectangle.

It has also been shown that the absence of the Gln172 side-chain reduces the affinity of the enzyme for its substrates, but increases the rates of catalysis. This residue potentially has a role in the stabilisation of the $\beta_2\alpha_2$ loop during catalysis and in its absence, the loss of interactions between Gln172′ and the backbone carbonyl of Lys99 and the peptide amino group of Arg101, may prime *Nme*DAH7PS for entry into its regulated state. This may explain the hypersensitivity of this variant to concentrations of Phe. This feature is also observed to

some degree in the *Nme*DAH7PS^{Glu176Ala} variant, which also displayed increased sensitivity to Phe at low concentrations.

Instead of breaking these connections and revealing the functional effects, which has been described in detail here, further work should aim to strengthen the identified interactions connecting allosteric and active sites. This would involve, changing hydrogen bonding residues into salt-bridge forming residues and may result in enzymes with the ability to bind Phe, but the inability to undergo the appropriate allosteric mechanisms. This may assist in the deconvolution of the intra and inter chain contributions to allosteric regulation by Phe.

It may also be useful to test the aforementioned variant enzymes across a wide range of pH values, in order to marry their allosteric sensitivity at varying pH to their proposed pK_a change observed from the pK_a calculations. Identification of the effects of Phe at a variety of concentrations on apparent K_m values for these variants may also give interesting insight into their contribution to the allosteric signal and may enable better understanding of the kinetics and binding isotherms of these variants. This would provide in the methodologies used to determine the communication pathway(s) in *Nme*DAH7PS.

6.9 Conclusion

The complicated interplay of residues in *Nme*DAH7PS required for catalysis and regulation are very challenging to deconvolute. The subtle mechanisms adopted by the enzyme to achieve the requirements of the cell revolve around a single catalytic loop. While identification of the importance of this loop is relatively trivial, a true understanding of its dynamic motion is difficult to ascertain. The next step in the *Nme*DAH7PS chapter will be to employ dynamic techniques to better understand the interplay between residues of this loop and surrounding residues, substrates and synthetic inhibitors. While it is very selective and system specific, the methodologies employed to understand this enzyme extend far beyond this and can be applied to any protein that is up- or down-regulated by its local environment.

It is truly fascinating that nature has evolved such a simple, yet intricate mechanism to achieve and regulate the flux of the shikimate pathway in *Nme*DAH7PS based on the cellular demand for metabolites.

CHAPTER 7- GENERAL EXPERIMENTAL DETAILS

7.1.1 Reaction conditions and work-up

All reactions were carried out in flame dried glassware under an inert atmosphere of nitrogen gas and stirred with magnetic stirrers unless otherwise stated. Removal of solvents *in vacuo* was carried out using a rotary evaporator under reduced pressure with the water bath not exceeding 40 °C unless otherwise stated. Solid products were dried using an oil pump vacuum. Reactions at -78 °C were cooled using a bath of solid carbon dioxide/acetone.

7.1.2 Solvents and reagents

Reagents were obtained from commercial suppliers and used without further purification. Before use, solvents were dried by passage through alumina columns in a solvent purification system custom built by the Chemistry Department workshop.

DOWEX® 50WX8-100 was used in a packed column (~1x10cm). Before use ~100 mL of 4M aqueous hydrochloric acid was passed through the column and then the column was brought to neutral pH with continuous washing with water. Thorough regeneration involved washing with hydrochloric acid as stated above, followed by 50-100 mL 1M aqueous sodium hydroxide solution. The sodium hydroxide solution was washed off the column with multiple column volumes of water until the eluate was near neutral pH. This was followed by another acid wash before use.

7.1.3 Chromatography

Merck Kieselgel 60F254 pre-coated aluminium-backed plates were used for thin layer chromatography. Visualisation of spots was achieved using a UV lamp (254 nm), and by staining with basic potassium permanganate solution or phosphomolybdic acid (PMA), followed by gentle heating with a hot air gun. Flash chromatography was carried out using Silica 60 230-400 mesh and analytical grade solvents.

7.1.4 NMR spectroscopy

NMR spectroscopy was carried out on an Agilent 400NMR and Varian VNMR 500 spectrometers. ^{1}H-NMR spectra were calibrated to the signal of the residual protons of the deuterated solvent stated and chemical shifts δH are reported relative to tetramethylsilane. ^{13}C-NMR spectra were calibrated to the signal of the deuterated solvent indicated and

chemical shifts δC are reported relative to tetramethyl silane. ^{31}P-NMR spectra were calibrated and chemical shifts δP were relative to an external standard of 85% phosphoric acid. All chemical shifts are reported in parts per million (ppm) and coupling constants (*J*) in Hertz (Hz).

7.1.5 Mass spectrometry

High resolution mass spectrometry was carried out on a Bruker maXis 3G or Micromass LCT Classic spectrometers, using positive or negative electrospray ionisation.

7.1.6 UV-visible spectroscopy

UV spectrophotometry was carried out on a Varian Cary® UV 100 spectrophotometer in stoppered quartz cuvettes of 2mm or 10mm path length. The standard assay temperature of 25 °C was maintained using a Varian Peltier thermostat connected to the jacketed multi-cell holder.

7.1.7 pH measurements

The pH of buffers and other aqueous solutions was determined using a UB10 Ultrabasic benchtop pH meter from Denver Instrument Company or a Seven Compact S220 pH meter from Mettler Toledo. For samples of small volume, a micro glass body combination pH electrode was used. The sample pH was adjusted using aqueous hydrochloric acid (1M or 10M) and sodium hydroxide (1M or 10M) solutions unless otherwise stated.

7.1.8 Preparation of buffers and kinetic assay solutions

All buffers and assay solutions were prepared using water from a Millipore Milli-Q system. This water is referred to as Milli-Q throughout this thesis. All buffers and assay solutions were filtered prior to use, after the addition and solvation of the last solid component.

The buffers used contained:

-Lysis Buffer:

10mM 1,3-*bis* (*tris* (hydroxymethyl methylamino) propane (BTP) pH 7.3, 1 mM ethylenediamine tetraacetic acid (EDTA), 200 µM phosphoenol pyruvate (PEP) and 200 mM potassium chloride (KCl).

-Anion exchange (Source-Q) and Hydrophobic interaction (Source-Phe) binding buffer:

10mM BTP pH 7.3, 1 mM EDTA and 200 µM PEP.

- Anion exchange (Source-Q) elution buffer:

10mM BTP pH 7.3, 1 mM EDTA, 200 µM PEP and 1 M sodium chloride (NaCl.

-Hydrophobic interaction (Source-Phe) elution buffer:

10mM BTP pH 7.3, 1 mM EDTA, 200 µM PEP and 1 M ammonium sulphate.

-Size exclusion chromatography (SEC) buffer:

10mM BTP pH 7.3, 200 µM PEP and 100 mM KCl.

-Assay buffer

50mM BTP pH 6.8

For the preparation of solutions used in kinetic assays, solid PEP monopotassium salt (Alfa Aesar and Sigma Aldrich) was dissolved in the appropriate volume of assay buffer to yield a 10mM stock solution. Manganese (II) sulphate solutions and inhibitor solutions were prepared by dissolution of the solid in Milli-Q water. All kinetic assay solutions were filtered through a 0.20 µM syringe filter before use.

7.1.8.1 Preparation of E4P

Solutions of erythrose 4-phosphate E4P for kinetic assays were prepared using a variation of previously described methods.[152,153] Glucose 6-phosphate (G6P, 547 mg, 1.94 mmol, monosodium salt) was dissolved in 3 mL water and 350 mL glacial acetic acid added. This solution was placed in a three-necked round-bottom flask under a nitrogen stream and equipped with a stirrer bar and dropping funnel. Lead tetraacetate (1.50 g, 3.40 mmol, 1.75 eq) was dissolved in

glacial acetic acid (75 mL) and concentrated sulfuric acid (0.9 mL) added to this solution. The lead tetraacetate solution was added to the G6P solution dropwise over 4 h at room temperature. The presence of excess oxidant should be avoided and can be checked by moistened starchiodide paper. The addition is accompanied by the formation of a white lead sulphate precipitate. The resulting reaction mixture was stirred for another hour at room temperature, filtered through celite and the filter cake washed with water (200 mL). The

combined acetic acid and aqueous filtrates were concentrated to about 50 mL using a rotary evaporator under oil-pump vacuum and equipped with a dry-ice/acetone condenser. The solution was then diluted with water (100–200 mL) and concentrated again, and repeated 3–5 times to remove acetic acid. Alternatively, removal of acetic acid was achieved using overnight continuous liquid-liquid extraction of the aqueous phase with ether. The resulting crude solution of E4P was purified using anion exchange chromatography on Source 15Q resin, eluting with a gradient of ammonium formate (0–2 M) in 0.2 M aqueous formic acid. E4P-containing fractions were identified by enzyme assay with *Nme*DAH7PS (see page 174) and pooled. The combined E4P-containing fractions were passed down a column packed with DOWEX 50WX8-100 resin (H+-form) and the eluate concentrated by rotary evaporation (see above) to yield the desired E4P-concentration (20–30 mM as determined by enzyme assay, see page 174). Solutions were aliquoted (500 µL), snap-frozen in liquid nitrogen and stored at −80 ºC. In order to allow for equilibration of monomeric, dimeric and oligomeric species of E4P, solutions were thawed the day before use, stored at 4 ºC overnight and equilibrated at room temperature for at least 1 hr before kinetic assays were carried out.

7.1.8.2 LANZETTA PHOSPHATE ASSAY

Lanzetta reagent[154] was prepared fresh as required from the following components: 3 parts 0.045% w/v malachite green in water, 1 part 4.2% w/v ammonium molybdate in 4 M HCl, 0.1 parts 1.5% v/v Triton X-100 in water. The components were mixed in the dark and stirred for 1 h before the solution was filtered through a 0.45 µm syringe filter. For the qualitative detection of phosphate-containing fractions after anion exchange chromatography, a 20 µL sample of each fraction was mixed with 250 µL of Lanzetta reagent and the colour change judged by optical inspection. For the quantitative determination of inhibitor concentration, 300 µL of the inhibitor solutions were incubated with 10 µL calf alkaline phosphatase solution (5 units/mL in 4 mM MgCl2) for at least 2 h. To 100 µL of the digested sample was added 700 µL Lanzetta reagent and the absorbance at 630 nm determined after 20 minutes. A calibration curve for the determination of phosphate concentration was obtained from analogous analysis of solutions of appropriate concentrations (6-150 µM) of KH2PO4, which had been dried in high vacuum for at least 3 h before use. As a control, a glucose-6-

phosphate solution of known concentration was also digested with calf alkaline phosphatase and analysed.

7.1.9 Removal of metal ions from solution

Divalent metal ions were removed from solutions by treatment with Chelex® 100 resin (Bio-Rad). The resin was added to a solution and then stirred for at least 1 h before the resin was removed by filtration (0.2 µm). The resin invariably caused the pH of the solution to increase, sometimes necessitating adjustment to correct the pH of any resin-treated solution.

7.1.10 Glycerol stocks

All plasmid-containing strains of *E. coli* were stored frozen at −80 °C as glycerol stocks. Glycerol stocks were created by mixing 0.4 mL of 50% (v/v) glycerol with 0.6 mL of an overnight culture in either a 1.6mL micro-centrifuge or screw-top tube.

7.1.11 Antibiotics

The genes for *Nme*DAH7PS had previously been cloned into the pT7-7 vector, which carries resistance to ampicillin. Ampicillin was added to all growth media to a final concentration of 0.1 mg/mL. Stock solutions of ampicillin at 100 mg/mL were created, filter sterilised and stored at −80 °C, and freeze-thawed a maximum of two times.

7.1.12 Site-directed mutagenesis

Mutagenesis was performed using Phusion® High-Fidelity DNA Polymerase (Thermo Fisher Scientific). PCRs were performed in an Veriti® 96-well Thermal Cycler (Applied Biosystems) or an iCycler (Bio-Rad). Typically each PCR was 20 µL and used 50 $ng\mu l^{-1}$ of plasmid template and 50 $ng\mu l^{-1}$ of each primer. Cycling

protocols were those recommended by each kit or polymerase. Primers for mutagenesis were synthesised by life technologies and resuspended in either TE buffer (10mM Tris-HCl, 0.1mM EDTA) or sterilised Milli-Q water to a concentration of 100 µM. Template DNA (methylated) was digested by treatment with the restriction enzyme Dpn1 according to the protocols of the kit.

7.1.13 Plasmid extraction and purification

Plasmids were extracted using a High Pure Plasmid Isolation Kit (Roche). Small cultures (5 mL) of plasmid-containing cells were grown overnight with the appropriate antibiotic(s). The cultures were harvested the following morning and plasmids extracted and purified as per the kit instructions, with two variations. Firstly, as optionally specified in the kit instructions, the glass fibres were pre-equilibrated with binding buffer to increase plasmid yield, and secondly, plasmids were eluted from the fibres using 70 µL to 100 µL rather than 100 µL of elution buffer, to ensure a sufficient plasmid concentration for downstream uses. The concentration of purified plasmid was measured by absorption at 260nm using an appropriately blanked Nanodrop ND-1000 spectrophotometer.

7.1.14 DNA Sequencing

All DNA sequencing was performed at the Massey Genome Servive (MGS) using an Applied Biosystems 3130xl Genetic Analyzer with BigDye Terminator v3.1 (Applied Biosystems) sequencing chemistry. Typically 400 ng to 500 ng of purified double-stranded plasmid and 4 pM of primer were supplied for each sample. As all DAH7PS genes were in the pT7-7 vector the same forward and reverse primers were used for all sequencing.

Forward: 5′ TAA TAC GAC TCA CTA TAG GGA GA
Reverse: 5′ GTT TAC TCA TAT ATA CTT TAG AT

7.1.15 Transformation

Chemically competent cells (50 µL to 100 µL aliquots) were thawed on ice before addition of 8 µL of either PCR product or purified plasmid. The mixture was left on ice for 20 min then placed in a 42 °C. water bath for 30 s to 45 s before being placed on ice again for 2 min to 3 min. Cells were then outgrown for 1 h at 37 °C while shaking after addition of 250 µL super optimal broth (SOC) medium. A portion (50 µL to 200 µL) of the transformed cells was spread directly onto an LB-agar plate containing the appropriate antibiotic(s) and left to grow on the inverted medium overnight at 37 °C.

SOC medium: 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgSO4, 20mM glucose. Filter sterilised in aliquots.

7.1.16 Cell lines

The cell lines used for plasmid propagation were either *E. coli* XL1-Blue or TOP10 (Invitrogen). For protein expression either *E. coli* BL21 (DE3) or BL21 (DE3) Star (Invitrogen) cell lines were used.

7.1.17 Culture media

LB-agar was prepared by dissolving either LB-agar (Miller's) (37 gL^{-1}) or LB (Lennox L) (20 gL^{-1}) and agar (15 gL^{-1}) base in Milli-Q water and sterilised by autoclaving. The LB-agar solution was then heated in a microwave oven until boiling, and left to cool in a 60 °C oven. The LB-agar solution was then further cooled to 50 °C before antibiotic(s) were added immediately prior to pouring into round petri dishes.

LB media for plasmid extraction, pre-cultures and protein expression cultures were prepared by dissolving 20 gL^{-1} LB (Lennox L) base in Milli-Q water and sterilised by autoclaving. Antibiotic(s) were added immediately prior to use.

7.1.18 Protein expression cultures

Cultures for protein expression were grown in 2 L baffled conical flasks in a shaking incubator at 180 rpm. Expression cultures were inoculated with an overnight culture ~ 1/40 the volume and grown at 37 °C until induction of protein expression.

Transcription of the DAH7PS genes were under the control of an inducible *lac* operon, therefore protein expression was induced by addition of IPTG to a final concentration of 1 mM, when cultures were at a mid-logarithmic phase (OD600 of 0.4AU to 0.6AU). After induction, cultures were transferred to 23 °C and harvested the following morning.

7.1.19 Cell harvesting

Large cultures were harvested in 1 L bottles at 12 000 *g* for 15 min at 4 °C. The pellets were then re-suspended in a small volume of the supernatant liquid and transferred to 50mL centrifuge tubes and centrifuged at 6000 *g* for 10 min at 4 °C. The supernatant liquid was

removed and the cell pellet either immediately lysed, or flash frozen in liquid nitrogen and stored at −80 °C.

Small volume cultures (≤ 5 mL) were progressively harvested in 1.6mL micro-centrifuge tubes at 16 000 *g* for 1 min.

7.1.20 Cell lysis

Cells were lysed by sonication. Sonication was performed with an Omni-Ruptor 4000 Ultrasonic Homogenizer (Omni International). Cell pellets were resuspended in chilled lysis buffer (typically 25 mL) on ice, and sonicated in a beaker surrounded by packed ice. They were sonicated at 80% power and 5 multiples of 5 min with 40-50 percent pulsation.

Cellular debris were removed by centrifugation either at 40 000 *g* for 30 min at 4 °C.

7.1.21 Protein purification

All proteins were purified by a three-step procedure composed of, in step order, anion-exchange, hydrophobic interaction and size exclusion chromatography. The chromatography was performed on a GE Healthcare AKTA purifier™ 10 machines.

Before use, all buffers were filtered (0.2 µm) under vacuum and degassed. Samples were all likewise filtered before being loaded to either a 10 mL,50 mL or 150 mL Superloop™ (GE Healthcare), or, for samples of volume less than 5 mL, into a static loop. The eluate was fractionated in 2mL fractions collected in 96-well plates. On A¨KTA purifier™ 10 machines, the eluate from injection of the sample onto the column (flow-through) was collected in 50 mL centrifuge tubes. Elution from the columns was monitored principally at 280 nm, and depending on the machine being used, also at 260nm and 214 nm. When the fractions corresponding to DAH7PS could not be deduced from the chromatogram, suspected DAH7PS-containing fractions were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and/or for DAH7PS activity. Fractions containing DAH7PS were then pooled and prepared for the next chromatographic step.

7.1.21.1 Anion exchange chromatography

AEC was performed using SOURCE™ 15Q resin packed in a Tricorn 10/100 column (GE Healthcare), at 4 °C. The supernatant liquid (resulting from removal of the cellular debris by

centrifugation after cell lysis) was diluted four-fold with a no-salt buffer to reduce the concentration of anions. The no-salt buffer was 10mM BTP at pH 7.3 (*Nme*DAH7PS, pI = 6.28), 200 µM PEP and 1mM EDTA. Elution from the column was performed using gradients of increasing NaCl concentration up to 1M at a flow rate of 1 $mLmin^{-1}$ to 3 $mLmin^{-1}$.

For *Nme*DAH7PS and mutants thereof, the elution protocol involved a linear gradient to 0.4 M NaCl over (0.8 x loaded volume). The gradient was then removed and 1 M NaCl was used to elute off any other bound protein.

7.1.21.2 Hydrophobic interaction chromatography

The HIC was performed using SOURCE™15Phe resin packed also in a Tricorn 10/100 column (GE Healthcare). This purification step was performed at room temperature, as at 4 °C the hydrophobic interactions are weaker and the resin performance is degraded. Elution was performed at a flow rate of 1 $mLmin^{-1}$ to 3$mLmin^{-1}$.

Crystalline $(NH_4)_2SO_4$ was added to pooled AEC fractions to a concentration of 1M before the sample was filtered and loaded into a loop. The same buffers as used in anion exchange were used for HIC, except $(NH_4)_2SO_4$ was substituted in place of NaCl. For all proteins, elution was by linear gradient from 1M to 0M $(NH_4)_2SO_4$ over 50 mL. *Nme*DAH7PS and mutants eluted, depending on flow rates and column condition, between 0.75M and 0.25M $(NH_4)_2SO_4$.

7.1.21.3 Size exclusion chromatography

Fractions containing *Nme*DAH7PS from HIC were pooled and loaded into a loop. SEC was performed using a HiLoad 26/60 Superdex™ 200 prep grade column (GE Healthcare), at 4., in the same no-salt buffer used in the both AEC and HIC. The elution flow rate was 1.0 $mLmin^{-1}$.

7.1.22 Polyacrylamide gel electrophoresis

SDS-PAGE was performed using NuPAGE® 10% Bis-Tris Gel 1.0mm × 12-well pre-cast gels (life technologies) in NuPAGE® MOPS SDS Running Buffer (life technologies). Electrophoresis was performed using a XCell SureLock™ Electrophoresis Cell (life technologies), also at 165V for 45 min. Samples were mixed with NuPAGE® LDS Sample Buffer (4×) and a reducing agent (DTT).

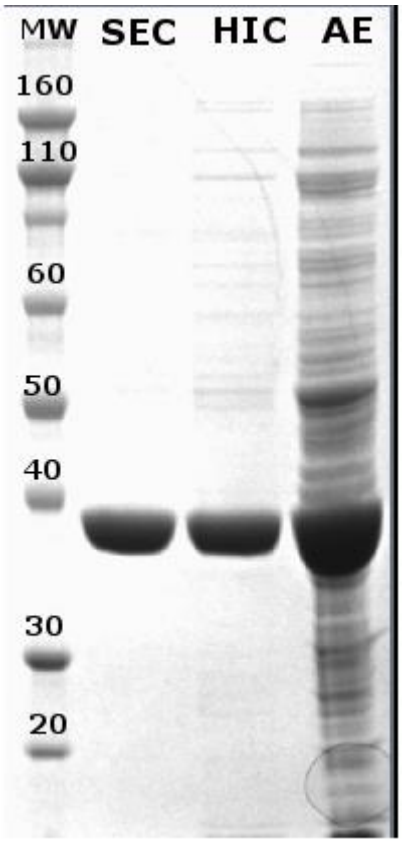


Figure 7.1.22.1. Typical purification of NmeDAH7PS^{WT} and variants, indicating the three steps of purification and the purified protein from the SEC.

7.1.23 Visualisation of protein bands

All SDS-PAGE gels were stained in square petri dishes, mostly by coomassie brilliant blue R-250, for at least 20 min before being destained to remove excess dye. Staining and destaining times were reduced by heating gel and solution in a microwave oven until almost boiling. Alternatively, some gels were stained using SimplyBlue™ Safe Stain (Invitrogen) instead of coomassie brilliant blue R-250. In the latter case, gels were first washed and also destained in water, and all steps were performed with heating until almost boiling in a microwave oven.

Stain: 0.1% (w/v) coomassie brilliant blue R-250, 10% (v/v) acetic acid and 40% (v/v) methanol

Destain: 10% (v/v) acetic acid and 40% (v/v) methanol

Molecular weight standards (Novex® Sharp Pre-Stained Protein Standards [life technologies]) were run as the sample in one lane of each gel. Photographs of gels were taken on a white light transilluminator in a Molecular Imager® Gel Doc™ XR (Bio-Rad).

7.1.24 Concentration and buffer exchanging of protein solutions

Protein solutions were concentrated using 10 000 Da molecular weight cut-off (MWCO) devices (Vivaspin 2 [GE Healthcare], Vivaspin 500 and Vivaspin 20 [Sartorius Stedim Biotech] and Amicon® Ultra-4 [Millipore]). All filtration units were rinsed before use with Milli-Q water. Protein solutions were buffer exchanged (e.g. into binding buffer for ITC) by repeatedly concentrating in a MWCO device and diluting with desired buffer.

7.1.25 Enzyme Storage

Purified protein was divided into aliquots in 0.2 µL thin-walled PCR tubes in 50 µL, flash-frozen in liquid nitrogen, and stored at −80 °C. All protein samples were rapidly thawed immediately before use and kept on ice. Protein solutions were stored at 10 mg/mL.

7.1.26 Protein concentration determination

The concentration of all protein solutions were measured by absorption at 280nm by an appropriately blanked Nanodrop ND-1000 spectrophotometer, using 2 µL samples of purified protein solution. The concentration was calculated from the absorption using molar extinction coefficient value for *Nme*DAH7PS is 31650 $M^{-1}cm^{-1}$. The coefficient values, calculated from protein sequences using the ProtParam tool on the ExPASy Proteomics Server,[111] were the same for both *Nme*DAH7PS^{WT} and variants. All measurements of protein concentration were performed as duplicates and averaged.

7.1.27 Standard enzyme activity assay

The standard assay for DAH7PS activity monitored the consumption of PEP by loss of absorbance at 232 nm, based on the method of Schoner and Herrmann[57]. All kinetic measurements were made using a Varian Cary 100 UV-visible spectrophotometer, stoppered 1 cm pathlength quartz cuvettes and a total assay volume of 1 mL. Standard assays were carried out at 25 °C and all assay solutions were equilibrated to temperature and baseline absorbance before initialisation with either enzyme, metal ion, or E4P. Standard reaction mixtures contained 50mM BTP pH 6.8, a variable concentration of PEP, E4P and enzyme, and when appropriate 100 µM MnSO4.

Assay buffers were treated with Chelex® 100 resin (Bio-Rad) prior to use to remove metal ions, and Chelex-treated buffer solution was used to create concentrated stock solutions of PEP and A5P. Metal-ion solutions were created using either Chelex-treated buffer or Milli-Q water.

Initial rates of reaction were measured as a least-squares fit of the initial rate data using Cary WinUV Kinetics Application (version 3.00, Varian). A unit of enzyme activity was defined as the loss of 1 µmol min^{-1} of PEP at 25 °C, and could be calculated from the measured loss of absorbance (at 232 nm) using Beer's Law (ϵ is $2.8\times10^{3} M^{-1} cm^{-1}$). The values of the kinetic parameters were determined by fitting data to the Michaelis-Menten equation using the software GraFit (version 5.0.13, Erithacus Software Limited).

7.1.28 Determination of substrate concentrations

The concentrations of substrate (e.g. PEP and E4P) solutions were determined using the standard assay system for DAH7PS activity. The concentration of the substrate not being measured was used in excess and the substrate for which the concentration was being measured was used at a limiting concentration. That the correct substrate was limiting was checked by addition of more of it after the reaction had reached completion. The change in absorbance, between that before initiation of the reaction and that at completion, was measured as ΔA_1. A control reaction, absent in one of the substrates, was used to measure the change in absorbance due to the addition of enzyme to the cuvette at initiation of the reaction (ΔA_2). The total change in absorbance ($\Delta A_1 + \Delta A_2$), was converted using Beer's Law to a concentration of the limiting substrate in the cuvette, and the stock substrate concentration calculated.

7.1.29 Isothermal titration calorimetry

All ITC experiments were performed on a VP-ITC microcalorimeter (Microcal, GE Healthcare) operating at 298 K. All experiments consisted of 29 or 56 injections: one 2 µL injection and 28 or 55 subsequent 10 or 5 µL injections (unless otherwise stated), using a reference power of 10 µCal s^{-1}. For every experiment the ligand was in the syringe and protein in the cell. Before use, protein solutions were buffer exchanged into a binding buffer consisting of 50mM BTP pH 7.0 and 0.5mM $MnSO_4$ for all *Nme*DAH7PS and mutants. The protein

concentration was measured by UV absorption and all solutions were degassed in a vacuum immediately before use and both the cell and syringe were washed several times with binding buffer.

The initial datum point was routinely deleted to allow for diffusion of ligand across the needle tip during the equilibration period. Heats of dilution experiments were measured independently and subtracted from the integrated data before curve fitting in Origin (version 7.0, OriginLab®) with the standard one site or sequential binding site model supplied by MicroCal.

7.1.30 Crystallisation of *Nme*DAH7PS and variants

Crystallisation of *Nme*DAH7PS was achieved by concentrating the protein to approximately 10 mg mL^{-1}.1 µL of the buffer-exchanged enzyme solution (9–11 mg mL^{-1}) was mixed with 1 µL of crystallisation buffer containing 0.1M Tris HCl (pH 7.2-7.5), 0.2 M trimethyl-amino-*N*-oxide (TMAO), 0.4 mM MnSO4 and 13–20% (w/v) PEG 2000MME. Crystals were grown by hanging-drop vapour diffusion over 500 µL of crystallisation buffer and the crystallisation trays were incubated at 20 °C. Crystals began to form in 48 hrs and were fully formed within three-seven days. Crystals were flash frozen using liquid nitrogen in a cryo-protectant solution containing reservoir solution and 20% (v/v) PEG400.

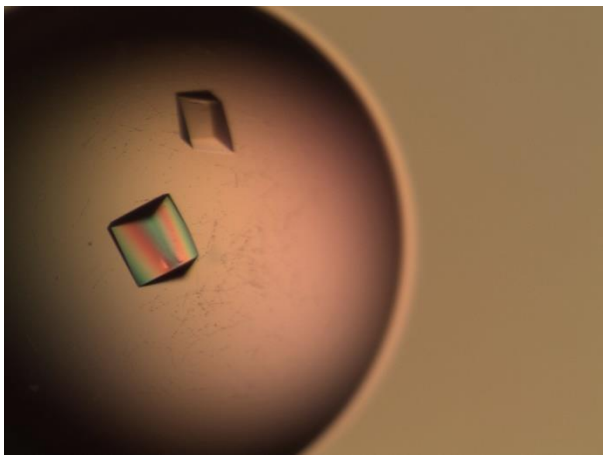


Figure 7.1.30.1 Typical crystals of NmeDAH7PS^{WT} used for soaking and seeding.

7.1.31 Seeding variant enzymes from *Nme*DAH7PS^{WT}

The growth of some variant enzymes was aided by the use of micro-seeding, if they would not grow in the conditions above. The use of fully grown crystals of *Nme*DAH7PS^{WT} were used as seeds for the growth of some, more difficult to crystallise variants. In this process, several hanging drops containing 1 µL crystallisation condition and 1 µL enzyme were placed on a crystal tray coverslip and left for one day. After allowing time for equilibration a grown *Nme*DAH7PS^{WT} crystal was crushed. This pool of microcrystalline *Nme*DAH7PS^{WT} was transferred into a corresponding clear drop of variant enzyme in a dilution series across the several clear drops to initiate crystal growth. This was achieved using a seeding tool Hampton Research™.

7.1.32 Co-crystallisation of active site inhibitors with *Nme*DAH7PS

Co-crystallisation of ligand with *Nme*DAH7PS was achieved by preparing a solution containing approximately 5mM inhibitor dissolved in 10mM BTP (pH 7.3). 400 µL of this solution was added to the stock solution containing 50 µL of 10 mg/mL enzyme and was concentrated to 50 µL using a 10 kDa molecular weight cut-off membrane (Vivaspin). The resulting solution was diluted again with 400 µL of inhibitor solution and concentrated by centrifugation. This process was repeated four times. The resulting protein solution was concentrated to approximately 10 mg mL^{-1}.

7.1.33 Soaking of ligands into *Nme*DAH7PS

Crystal soaking was achieved via two different methods. In the first method, crystals that were intended to be soaked in solutions containing ligands, multiple hanging droplets were prepared on the cover slide and equilibrated against reservoir solution in the same well. Droplets for crystallisation of *Nme*DAH7PS were setup using the method described above. The droplets for soaking were generated by mixing 1 µL of SEC buffer with 1 µL reservoir solution to which the desired concentration of the ligand/combination of ligands had been added. After crystals had formed and all droplets were deemed equilibrated against the reservoir solution (3-7 days), *Nme*DAH7PS crystals were transferred from the crystallisation drop to the soaking drop and left to soak for the time indicated.

In the other method, the crystallisation condition was brought to a ligand concentration of between 10 and 30mM. Upon formation of crystals, 1 µL of this solution was added to the 2 µL that contained the crystals. This brought the concentration of ligand in the drop to between 3 and 10 mM. This was left to soak for the indicated time.

7.1.34 Data collection, processing, structure determination and refinement

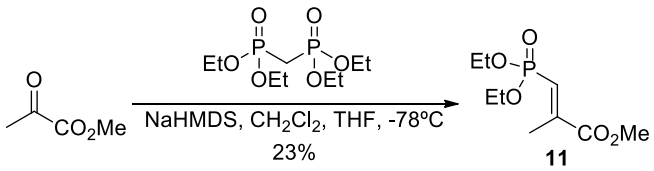
X-ray diffraction datasets were collected at the Australian synchrotron using the MX1 and MX2 beamlines[112]. The datasets were integrated and processed using iMosflm, XDS, Aimless (CCP4 program suite).[114113155156] The appropriate cut-off resolution was determined via CC½ ≥ 0.5 while ensuring the data set was complete in the highest resolution shell.[115] Matthews was run to determine number of chains present in the unit cell.[157,158] To minimise bias, all ligands and waters were removed from the search model (PDB code 4HSN) before molecular replacement (Phaser MR (CCP4)) was carried out.[116] Rigid body refinement in Refmac5 was used to generate the electron density maps for all structures and these were manually analysed to develop the structural model by means of COOT. [117,118] All datasets used the same sets of reflections for the calculation of R_{free} as were allocated in the search model (PDB code: 4HSN). The quality of the model was optimised by consecutive model building in COOT and refinement with Refmac5. Water molecules and ligands other than the ligands of interest were added to the structure, apart from regions of the structure, which showed evidence for the ligands of interest were added manually via interpretation of the (2|*Fo* - *Fc*|) and (|*Fo* - *Fc*|) maps, with the requirement that putative water molecules had at least one hydrogen-bonding partner. Residues Met1-Asp9 were also absent from the model (PDB code 4HSN) used for molecular replacement. Molprobity and the validation tools of COOT were used to assess structure quality during refinement cycles.[120] After the model and electron density was judged to be in good agreement in all regions apart from the ligand binding site(s) of interest, the electron density in the ligand binding site(s) was examined. If appropriate electron density was present, the ligand of interest, waters and potential other molecules were added to the model and the final rounds of refinement carried out. (2|*Fo* - *Fc*|) and (|*Fo* - *Fc*|) omit maps shown in figures were generated using the program FFT (CCP4 program suite) before addition of the ligands to the model.

7.1.35 Protein structure figures

Figures containing protein structures were generated using PyMOL (Version 1.7, Schrödinger LLC).

7.2.1 Synthesis

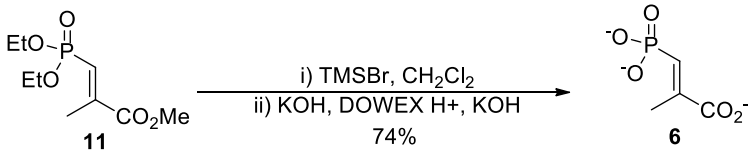
(*E*)-Methyl 3-(diethoxyphosphonyl)-2-methylacrylate



To a solution of tetraethyl methylene bisphosphonate (1.64 mL, 6.60 mmol) in THF (8 mL) at -78 °C under a stream of N_2 was added a solution of sodium bis(trimethylsilyl)amide in THF (7.2 mL, 6.10 mmol), the mixture was stirred for 30 minutes and allowed to heat to -35 °C over this time. The cold bath was replaced with an ice bath and methyl pyruvate (580 µL, 6.2 mmol) was added in a dropwise manner. The mixture was left to stir for 2.5 hours, and then quenched with saturated potassium dihydrogen phosphate (10 mL). The solution was concentrated *in vacuo* to remove the THF. The residue was extracted using EtOAc (5x10 ml). The combined organic layers were washed with brine and separated again. The washed organic layer was dried over $MgSO_4$. Purification by flash chromatography (50% EtOAc/PET ether) gave (*E*)-methyl 3-(diethoxyphosphonyl)-2-methylacrylate **11** (330 mg, 23%) and (*Z*)-methyl 3-(diethoxyphosphonyl)-2-methylacrylate (270 mg, 18.5%). NMR data were in accordance with literature values.[23]

R_f (50% EA/Pet Ether): 0.22. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 6.66 (dq, *J* = 16.0, 1.2 Hz, 1H), 4.09 (dq, J=7.8, 7.1 Hz, 4H), 3.79 (s, 3H), 2.25 (dd, *J* = 1.4, 3.3 Hz, 3H), 1.32 (t, *J* = 7.0 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 166.8 (d, *J* = 29.3 Hz, C_1), 146.9 (d, *J* = 9.0 Hz, C_2), 125.2 (d, *J* = 186 Hz, C_3), 62.0 (d, *J* = 5.6 Hz, $C_3PO\underline{C}H_2CH_3$), 52.6, 16.3 (d, *J* = 6.4 Hz, $C_3POCH_2\underline{C}H_3$), 15.8 (d, *J* = 6.4 Hz, C_4). ^{31}P NMR (162 MHz, $CDCl_3$, ^{1}H decoupled) δ ppm 15.0. ESI-MS positive ion (MH^+): Required 237.0886; found 237.0884.

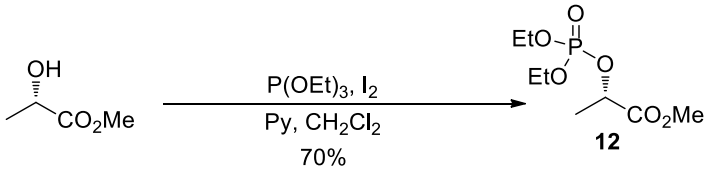
(*E*)-2-Methyl-3-phosphonatoacrylate (tripotassium salt)



To a solution of (*E*)-methyl 3-(diethoxyphosphonyl)-2-methylacrylate **11** (150 mg, 0.635 mmol) in DCM (5 mL) under a stream of N_2 was added trimethylsilyl bromide (600 µL, 4.5 mmol). The mixture was stirred overnight and then evaporated. The residue was dissolved in 10 % (w/v) aqueous KOH (2 mL), stirred for 5 minutes and then extracted with DCM (3x5 ml), and the organic phases discarded. The aqueous phase was treated with DOWEX-50X8 H^+-form resin, neutralised with 1M KOH and lyophilised to give a white powder. Analysis of the product by NMR showed only the presence of the desired (*E*)-2-methyl-3-phosphonatoacrylate **6**, and comparison of ^{31}P NMR signal integrals indicated that the powder was 35 % pure, the impurity being presumably inorganic potassium salts and/or waters of hydration, indicating a yield for this step of 74 % and an overall yield of 17 %. NMR data were in accordance with literature values.[23]

^1H NMR (400 MHz, D_2O) δ ppm 6.23 (dd, J = 15.6, 1.2 Hz, 1H), 1.95 (dd, *J* = 3.1, 1.2 Hz, 3H). ^{31}P NMR (162 MHz, D_2O, ^1H coupled) δ ppm 10.94. (dd, *J* = 17.2, 3.7 Hz). ^{13}C NMR (101 MHz, D_2O) δ ppm 176.4 (d, *J* = 24 Hz, C_1), 147.9 (d, *J* = 5.3 Hz, C_2), 126.3 (d, *J* = 173 Hz, C_3), 16.2 (d, *J* = 6.9 Hz, C_4). ^{31}P NMR (162 MHz, D_2O, ^1H decoupled) δ ppm 10.96. ESI-MS positive ion (MH^+): Required 167.0104; found 167.0104.

(*S*)-Methyl 2-(diethoxyphosphoryloxy)propanoate

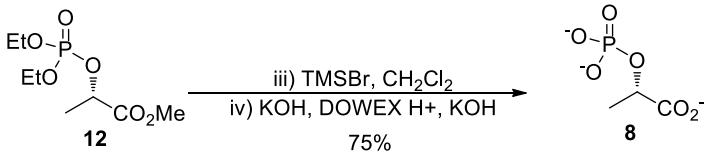


A solution of triethylphosphite (0.870 mL, 5.07 mmol) in DCM (12 mL) under nitrogen was cooled to 0 °C, and iodine (1.17 g, 4.61 mmol) was added. The solution was stirred until the characteristic brown iodine colour had been discharged, typically taking approx. 15 minutes. The resulting solution of diethylphosphoryl iodide was added drop-wise over 15 minutes to (*S*)-methyl lactate (367 µL, 3.843 mmol) in pyridine (2.5 mL) and DCM (60 mL) at 0 °C under nitrogen. Once addition was complete the mixture was stirred for 1 hour. The reaction was quenched with saturated aqueous sodium bicarbonate solution (20 mL), and stirred vigorously for 5 minutes. The organic phase was separated, and washed with 1 M aqueous hydrochloric acid (2 x 75 mL), saturated aqueous sodium bicarbonate solution (75 mL), and

brine (75 mL). The organic phase was dried over $MgSO_4$ and concentrated *in vacuo*. Flash chromatography (50% EtOAc/PET ether) gave the (*S*)-methyl 2-(diethoxyphosphoryloxy)propanoate **12** (640 mg, 70%). NMR data were in accordance with literature values.[23]

R_f (50% ethylacetate/hexane): 0.32. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 4.92 (qd, *J* = 7.1 Hz, 1H) 4.23 - 4.06 (m, 4H), 3.77 (s, 3H), 1.55 (d, *J* = 6.7 Hz, 3 H), 1.34 (dt, *J* = 3.1, 7.0 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ = 171.0 (d, *J* = 5.2 Hz, C_1), 71.6 (d, *J* = 5.2 Hz, C_2), 64.2 (d, *J* = 6.0 Hz, PO<u>C</u>$H_2$$CH_3$), 52.4 (d, *J* = 6.0 Hz, CO_2<u>Me</u>), 19.2 (d, *J* = 5.4 Hz, C_3), 16.0 (d, *J* = 5.4 Hz, $POCH_2$<u>C</u>H_3). ^{31}P NMR (162 MHz, $CDCl_3$) δ ppm -1.98. ESI-MS positive ion matrix (H^+): Required 241.0836; found 241.0833. $[\alpha]_D^{20}$ = -18.3°.

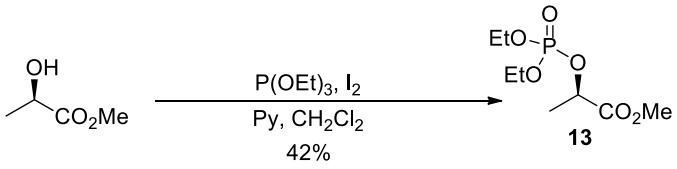
(*S*)-2-(Phosphoryl)propanoate (tripotassium salt)



To a solution of (*S*)-phosphoester **12** (462 mg, 1.91 mmol) in DCM (15 mL) under nitrogen was added trimethylsilyl bromide (1.54 mL, 11.5 mmol), and the mixture was stirred overnight then evaporated. The residue was dissolved in 10 % (w/v) aqueous potassium hydroxide (5 mL), stirred for 1.5 hours minutes then extracted with DCM (3 x 4 mL) and the organic phases discarded. The aqueous phase was treated with Dowex-50X8 H^+-form resin, neutralised with 1 M KOH and lyophilised to give a white powder, (1.03 g). Analysis of concentration via Lanzetta assay indicated that the (*S*)-phospholactate product **8** was 23 % pure containing ~20 % of lactate due to neighbouring group assisted hydrolysis. The remainder consisting of inorganic potassium salts and/or waters of hydration, giving a yield for this step of (237 mg, 75 %) and an overall yield of 52 % over two steps. NMR data were in accordance with literature values.[23]

^{1}H NMR (500 MHz, D_2O) δ ppm 4.34 (m, 1H, lactate), 4.27 (q, *J* = 1.3, 7.0 Hz, 1H), 1.6 (d, *J* = 7.0 Hz, 3H), 1.28 (d, *J* = 6.8 Hz, 3H, lactate). ^{13}C NMR (126 MHz, D_2O) δ ppm 180.5 (lactate), 178.2 (d, *J* = 5.3 Hz, C_1), 72.4 (d, *J* = 5.3 Hz, C_2), 47.7 (lactate), 22.5 (lactate), 19.9 (d, *J* = 3.8 Hz, C_3). ^{31}P NMR (162 MHz, D_2O) δ ppm 0.50 (80%); 0.10 (20%). ESI-MS negative ion, 20% AcOH matrix): Required 168.9907; found 168.9904.

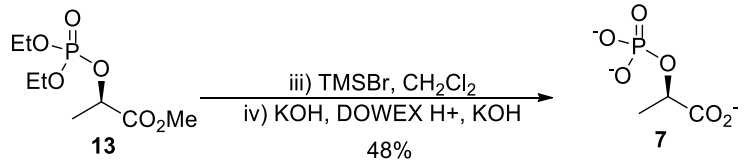
(*R*)-Methyl 2-(diethoxyphosphoryloxy)propanoate



A solution of triethylphosphite (0.870 mL, 5.07 mmol) in dichloromethane (12 mL) under nitrogen was cooled to 0°C, and iodine (1.17 g, 4.61 mmol) was added. The solution was stirred until the characteristic brown iodine colour had been discharged, typically taking approx. 15 minutes. The resulting solution of diethylphosphoryl iodide was added drop-wise over 15 minutes to (*R*)-methyl lactate (367 µl, 3.843 mmol) in pyridine (2.5 mL) and dichloromethane (60 mL) at 0°C under nitrogen. Once addition was complete the mixture was stirred for 1 hour. The reaction was quenched with saturated aqueous sodium bicarbonate solution (20 mL), and stirred vigorously for 5 minutes. The organic phase was separated, and washed with 1 M aqueous hydrochloric acid (2 x 75 mL), saturated aqueous sodium bicarbonate solution (75 mL), and brine (75 mL). The organic phase was dried over $MgSO_4$ and concentrated *in vacuo*. Flash chromatography (50% EtOAc/PET Ether) gave the phosphate ester **13** (380 mg, 42%).

R_f- (50%EA/Pet Ether)-0.32. ^{1}H NMR (400 MHz $CDCl_3$) δ ppm (qd, *J* = 7.1 Hz, 1H), 4.22 - 4.07 (m, 4 H), 3.77 (s, 3 H), 1.55 (d, *J* = 7.0 Hz, 3 H), 1.34 (dt, *J* = 3.3, 7.1 Hz, 6 H). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 171.0 (d, *J* = 5.2 Hz, C_1), 71.6 (d, *J* = 5.2 Hz, C_2), 64.1 (d, *J* = 6.0 Hz, $PO\underline{C}H_2CH_3$), 52.4 (d, *J* = 6.0 Hz, $CO_2\underline{Me}$), 19.2 (d, *J* = 5.4 Hz, C_3), 16.1 (d, *J* = 5.4 Hz, $POCH_2\underline{C}H_3$). ^{31}P NMR (162 MHz, $CDCl_3$) δ ppm -1.98. ESI-MS positive ion matrix (H^+): Required 241.0836; found 241.0833. $[\alpha]_D^{20}$ = +18.5°.

(*R*)-2-(phosphoryl)propanoate (tripotassium salt)



To a solution of (*R*)-phosphoester **13** (242 mg, 1.00 mmol) in DCM (7.7 mL) under nitrogen was added trimethylsilyl bromide (1.2 mL, 8.96 mmol), and the mixture was stirred overnight then evaporated. The residue was dissolved in 10 % (w/v) aqueous potassium hydroxide (3

ml), stirred for 1.5 hours minutes then extracted with DCM (3x 4 mL) and the organic phases discarded. The aqueous phase was treated with Dowex-50X8 H^+-form resin, neutralised with 1M KOH and lyophilised to give a white powder (470mg). Analysis of concentration via Lanzetta assay indicated that the (*R*)-phospholactate product **7** was 17 % pure and contained ~24 % lactate due to neighbouring group assisted hydrolysis. The remainder consisting of inorganic potassium salts and/or waters of hydration, giving a yield for this step of (80 mg, 48 %) and an overall yield of 20 % over two steps.

^{1}H NMR (400 MH, D_2O) δ ppm, 4.36 (m, 1H, lactate), 4.27 (d, *J* = 6.7 Hz, 1H), 1.62 (d, *J* = 7.0 Hz, 3H), 1.25 (d, *J* = 6.8 Hz, 3H, lactate). ^{13}C NMR (101 MHz, D_2O) δ ppm 180.5 (lactate), 178.2 (d, *J* = 5.3 Hz, C_1), 72.4 (d, *J* = 5.3 Hz, C_2), 47.7 (lactate), 22.5 (lactate), 19.9 (d, *J* = 3.8 Hz, C_3). ^{31}P NMR (162 MHz, D_2O) δ ppm 1.55 (76%); 1.29 (24%). ESI-MS negative ion, 20% AcOH matrix): Required 168.9907; found 168.9913.

The research described in chapter 2 was the subject of a publication before submission of this thesis. The supplementary material for this publication reporting experimental details appended with the permission of the copyright owners and can be accessed at:

SUPPLEMENTAL DATA

Structural analysis of substrate-mimicking inhibitors in complex with *Neisseria meningitidis* 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase - The importance of accommodating the active site water.

Logan C Heyes[1], Sebastian Reichau[1], Penelope J Cross[1], Geoffrey B Jameson[2] and Emily J Parker[1]

[1]Biomolecular Interaction Centre and Department of Chemistry, University of Canterbury, Christchurch, New Zealand

[2]Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand

Inhibition studies

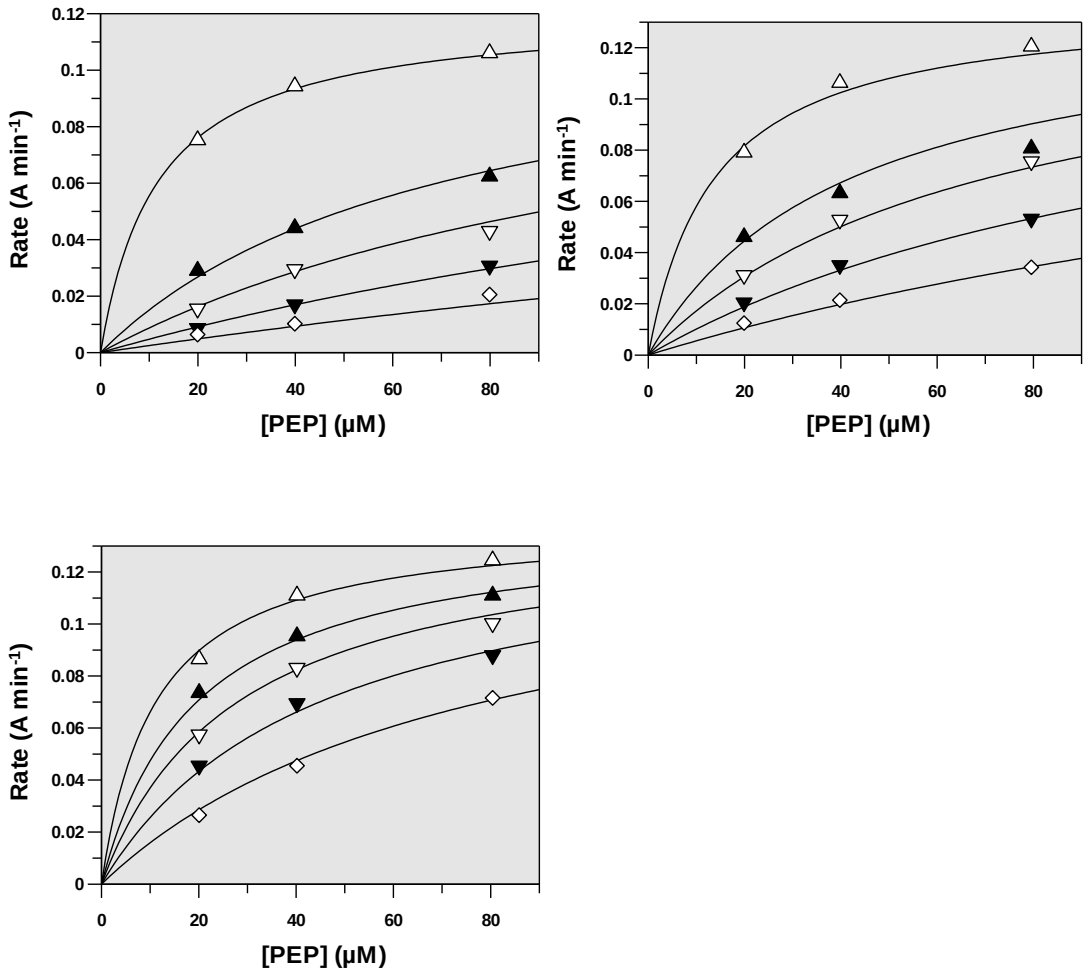
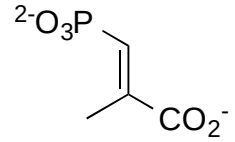
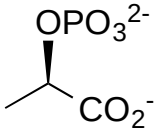
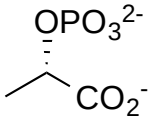


Figure S1. Inhibition of NmeDAH7PS., Top left. Inhibition with vinyl phosphonate 6 inhibitor using 19.9–79.6 µM PEP, 18.9–151.2 µM vinyl phosphonate inhibitor 6. Top right. Inhibition with (R)-phospholactate 7 using 19.9–79.6 µM PEP, 202–1616 µM (R)-phospholactate 7. Bottom left. Inhibition with of (S)-phospholactate 8 using 19.9–79.6 µM PEP, 268–2144 µM (S)-phospholactate 8. Enzyme activities were followed by monitoring the loss of PEP followed spectrophotometrically at 232 nm. Assay conditions used were 100 µM $MnSO_4$, 115 µM E4P. Initial rates were obtained by linear least squares regression of the absorbance data and the data fit is using non-linear regression analysis providing a global fit using all data points to the equation for competitive inhibition using Grafit (Erathicus software).

Table S1: Comparison of global fits to types of reversible inhibition

Inhibitor	Chi^2 values of fits to models for reversible inhibition[a]			
	Competitive	Noncompetitive	Mixed	Uncompetitive
 Vinyl phosphonate **6**	0.00004	0.0004	0.0035	0.0009
 (*R*)-Phospholactate **7**	0.0002	0.0005	0.0042	0.001
 (*S*)-Phospholactate **8**	0.00007	0.0004	0.008	0.0009

[a]The data was fitted to each equation using non-linear regression using Grafit. A graphical representation of the derived fits for competitive inhibition is proved in Figure S1. The following equations were used:

$$v = \frac{V_{max}[PEP]}{K_m\left(1+[I]/K_i\right)+[PEP]} \quad \text{(Competitive)}$$

$$v = \frac{V_{max}/\left(1+[I]/K_i\right)[PEP]}{K_m\left(1+[I]/K_i\right)+[PEP]} \quad \text{(Noncompetitive)}$$

$$v = \frac{V_{max}[PEP]}{K_m\left(1+[I]/K_i\right)+\left(1+[I]/K_i'\right)[PEP]} \quad \text{(Mixed)}$$

$$v = \frac{V_{max}[PEP]}{K_m+[PEP]\left(1+[I]/K_i\right)} \quad \text{(Uncompetitive)}$$

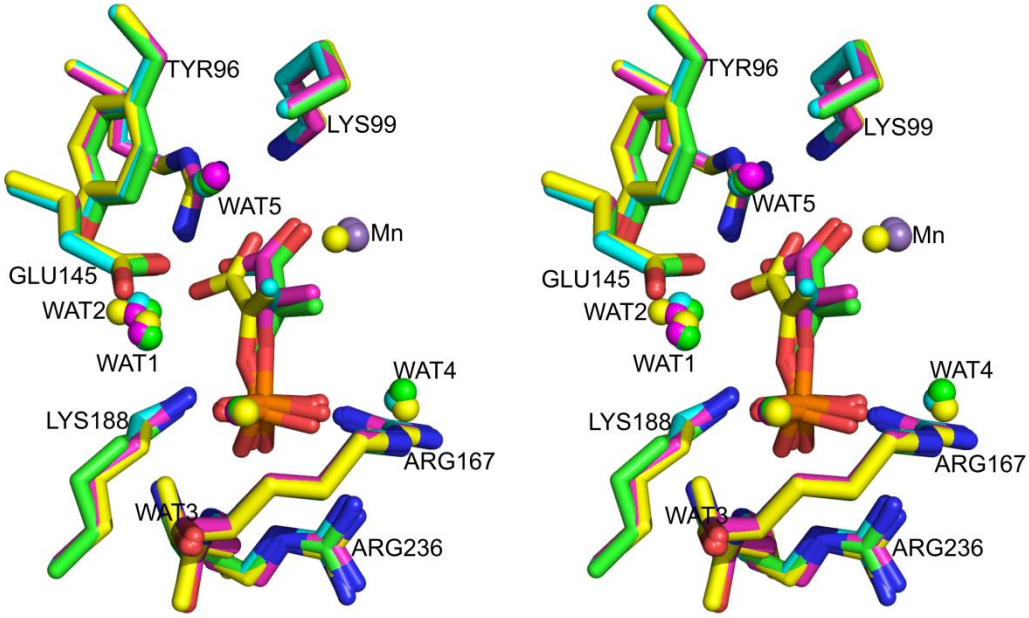


Figure S2. Stereoview of overlay of the three inhibitors and the natural substrate PEP. PEP 1 is shown in yellow, vinyl phosphonate 6 is shown in green, (R)-phospholactate 7 is shown in cyan and (S)-phospholactate 8 is shown in magenta. Mn^{2+} is shown as a purple sphere. The carbon atoms of the NmeDAH7PS residues and the key active site waters (WAT1-5) are shown in the same colours as the respective ligands.

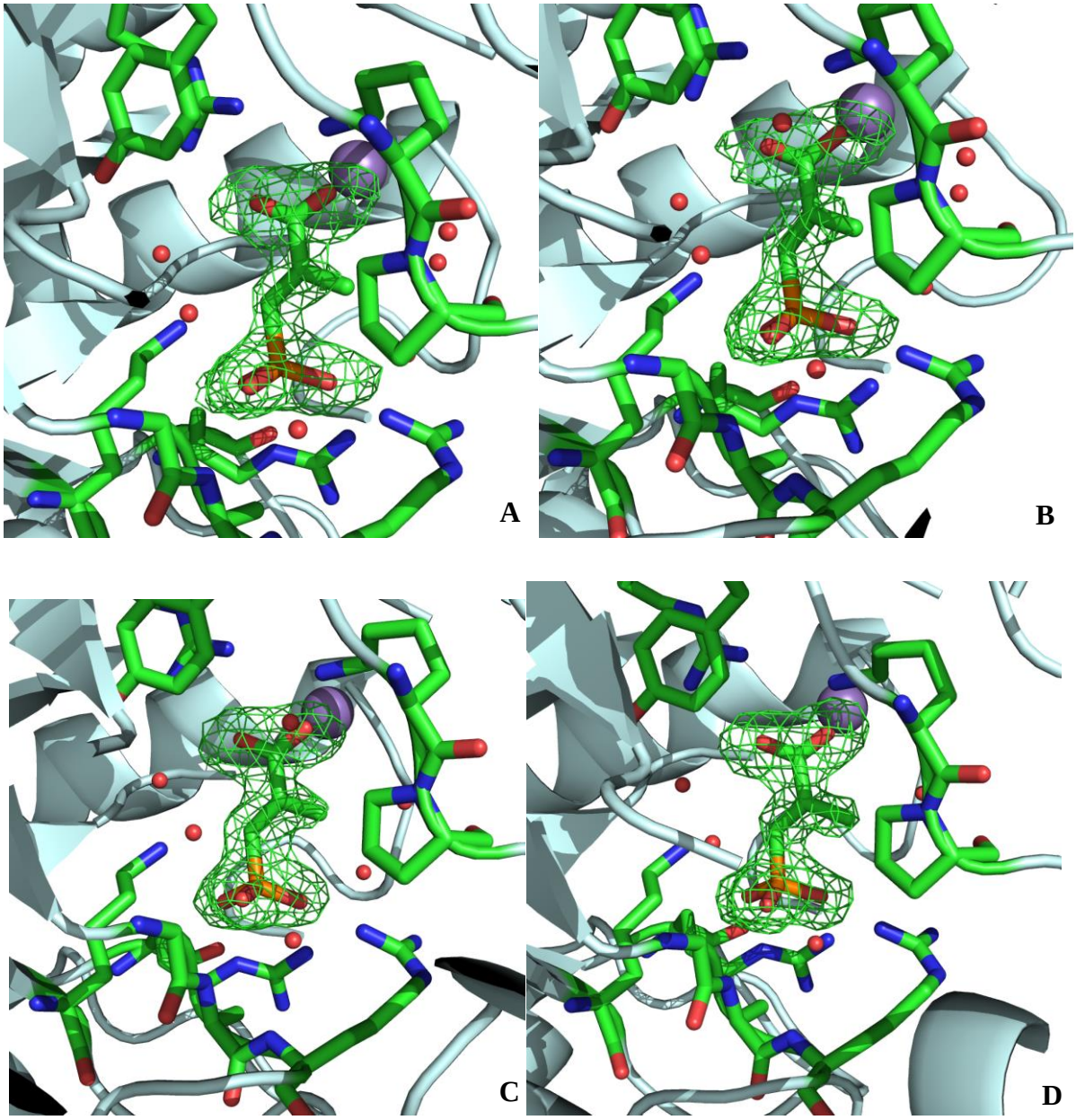


Figure S3. Omit maps of vinyl phosphonate 6 bound at the active site of NmeDAH7PS contoured at 3σ. The four chains of NmeDAH7PS are indicated. Top left: chain A, top right: chain B, bottom left: chain C and bottom right: chain D.

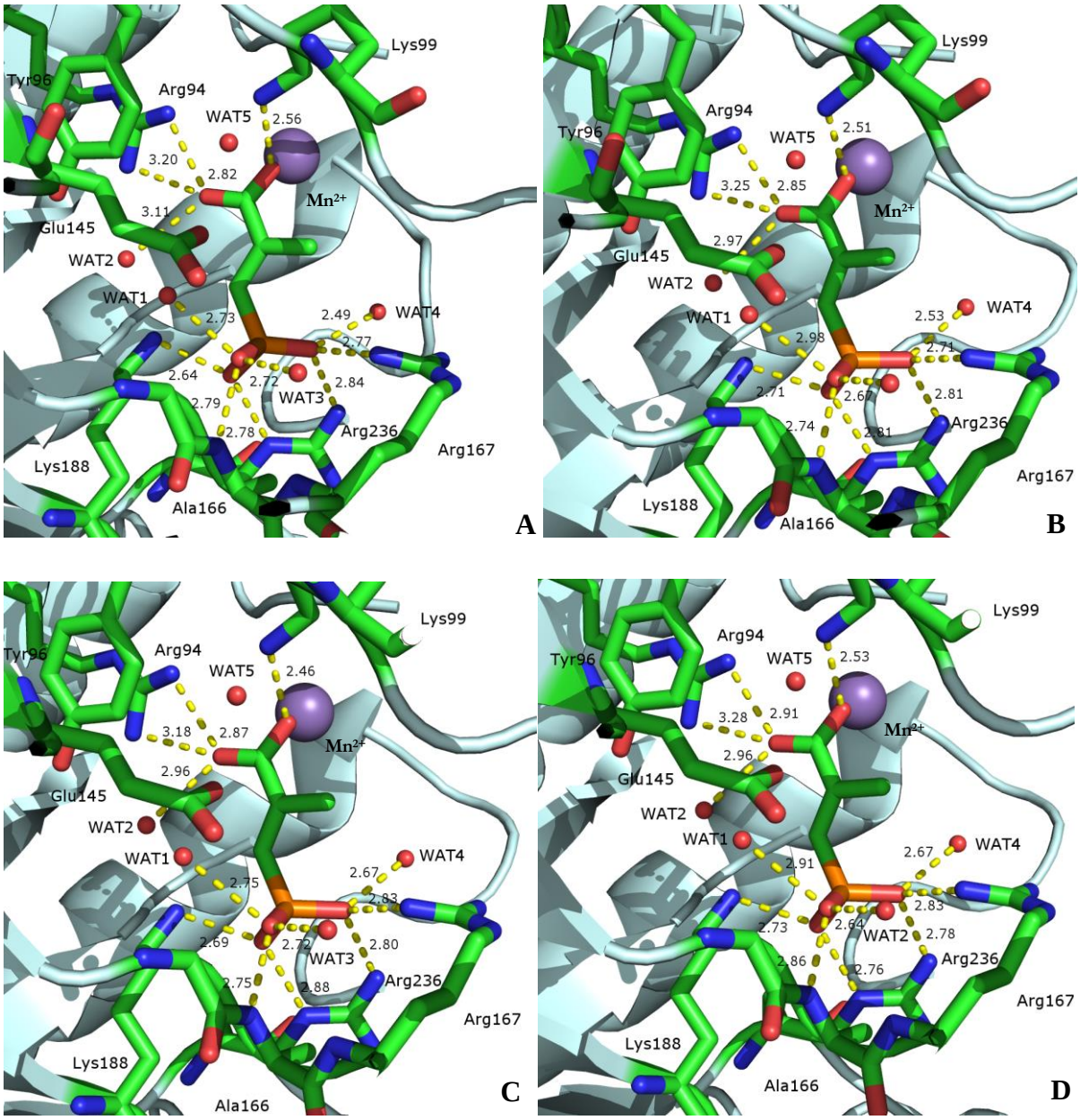


Figure S4. Interatomic distances of vinyl phosphonate 6 bound at the active site of NmeDAH7PS. The four chains of NmeDAH7PS are indicated. Top left: chain A, top right: chain B, bottom left: chain C and bottom right: chain D. Mn^{2+} is shown as a purple sphere in all panels.

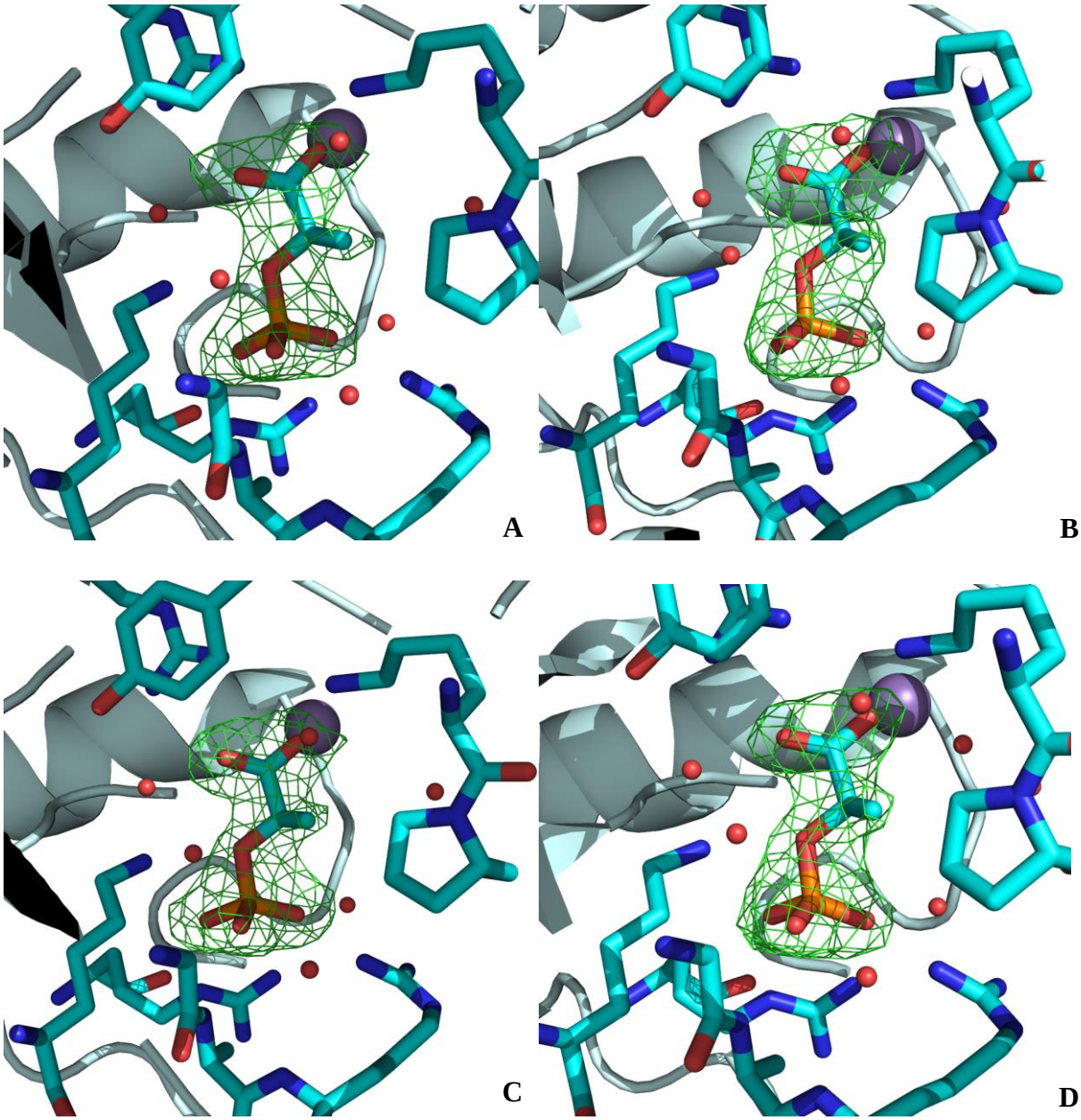


Figure S5. Omit maps of (R)-phospholactate 7 bound at the active site of NmeDAH7PS contoured at 3σ. The four chains of NmeDAH7PS are indicated: top left: chain A, top right: chain B, bottom left: chain C and bottom right: chain D.

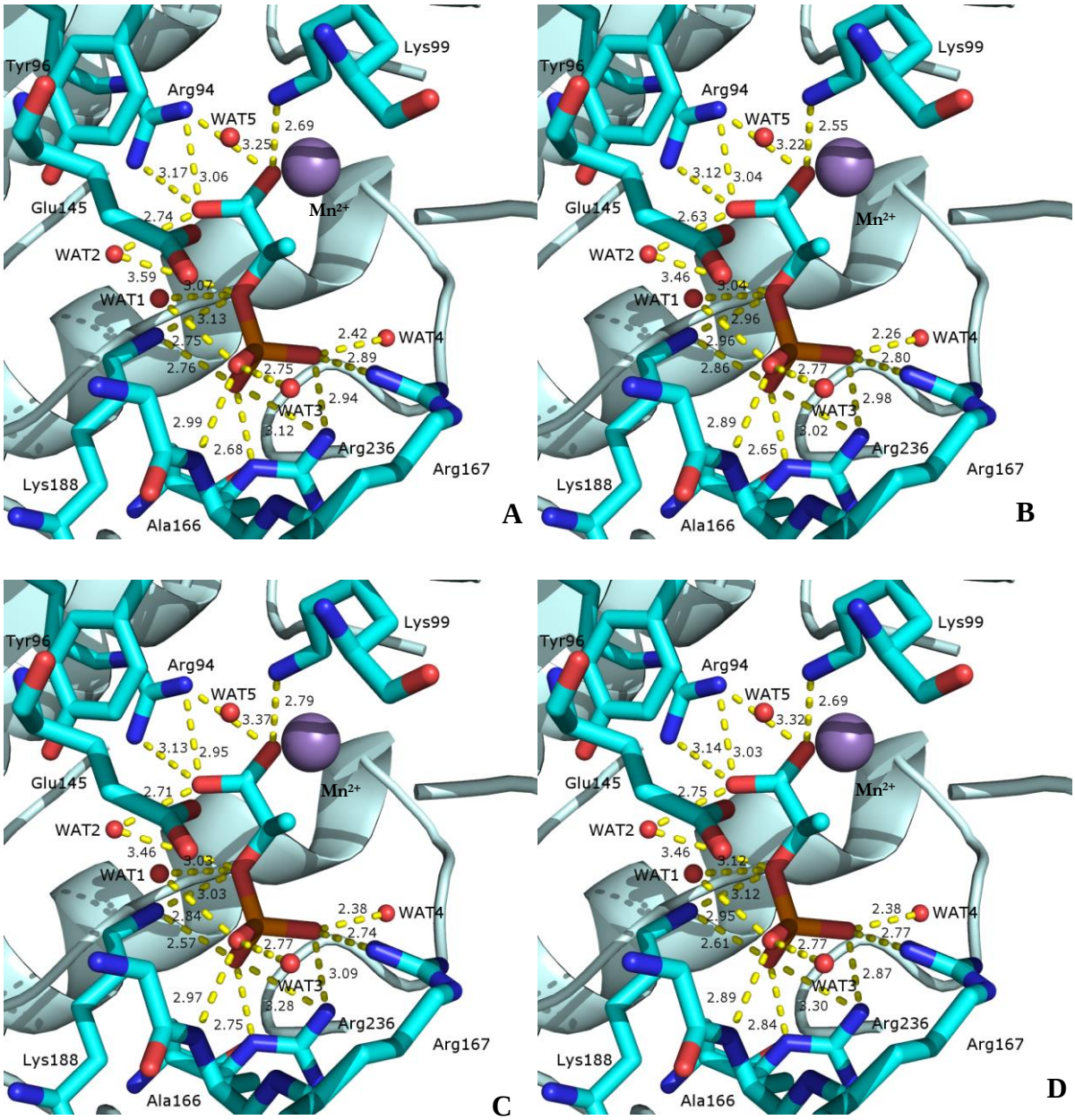


Figure S6. Interatomic distances of (R)-phospholactate 7 bound at the active site of NmeDAH7PS. The four chains of NmeDAH7PS are indicated: top left: chain A, top right: chain B, bottom left: chain C and bottom right: chain D. Mn^{2+} is shown as a purple sphere in all panels.

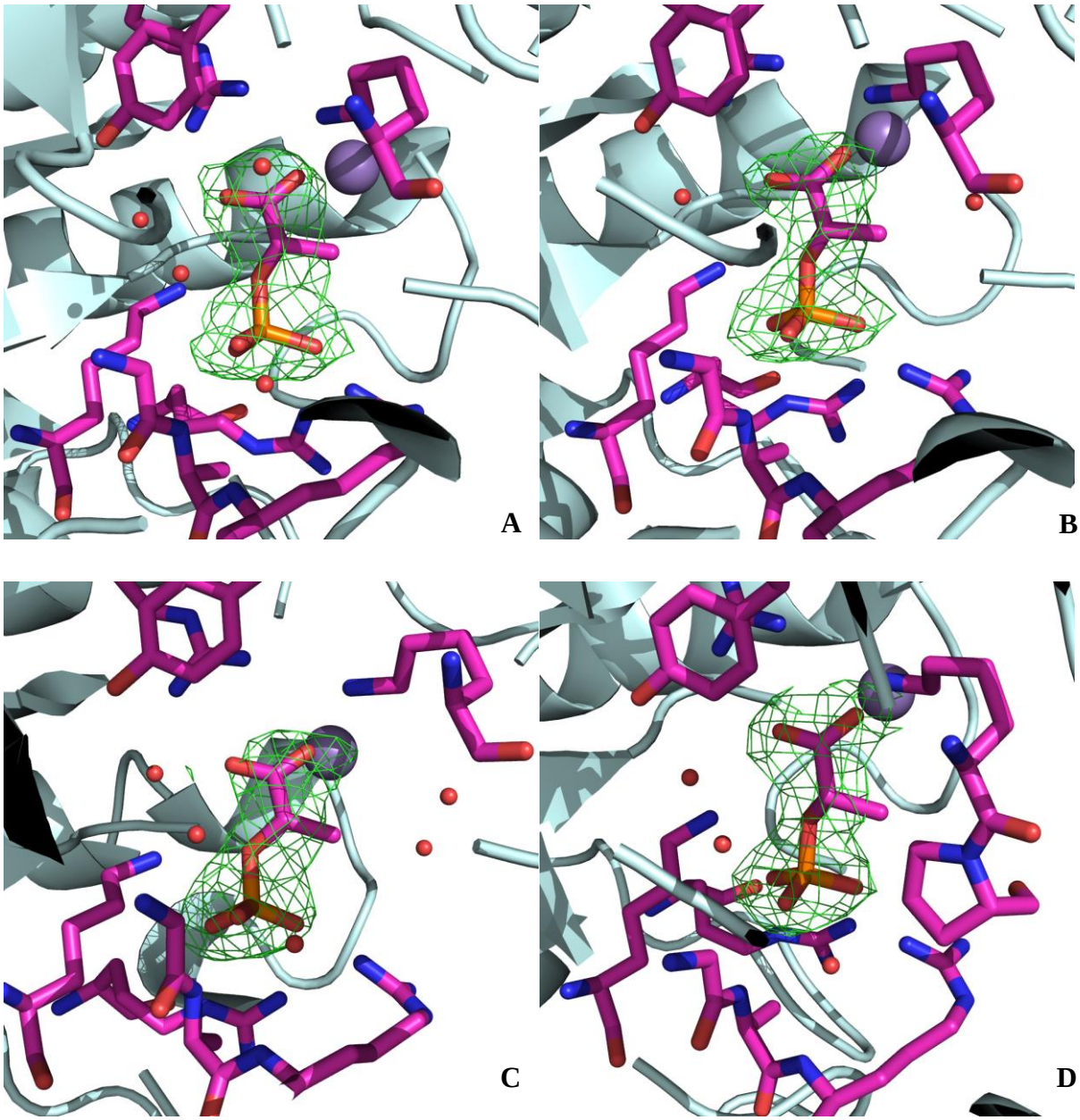


Figure S7. Omit maps of (S)-phospholactate 8 bound at the active site of NmeDAH7PS contoured at 3σ. The four chains of NmeDAH7PS are indicated. Top left: chain A, top right: chain B, bottom left: chain C and bottom right: chain D.

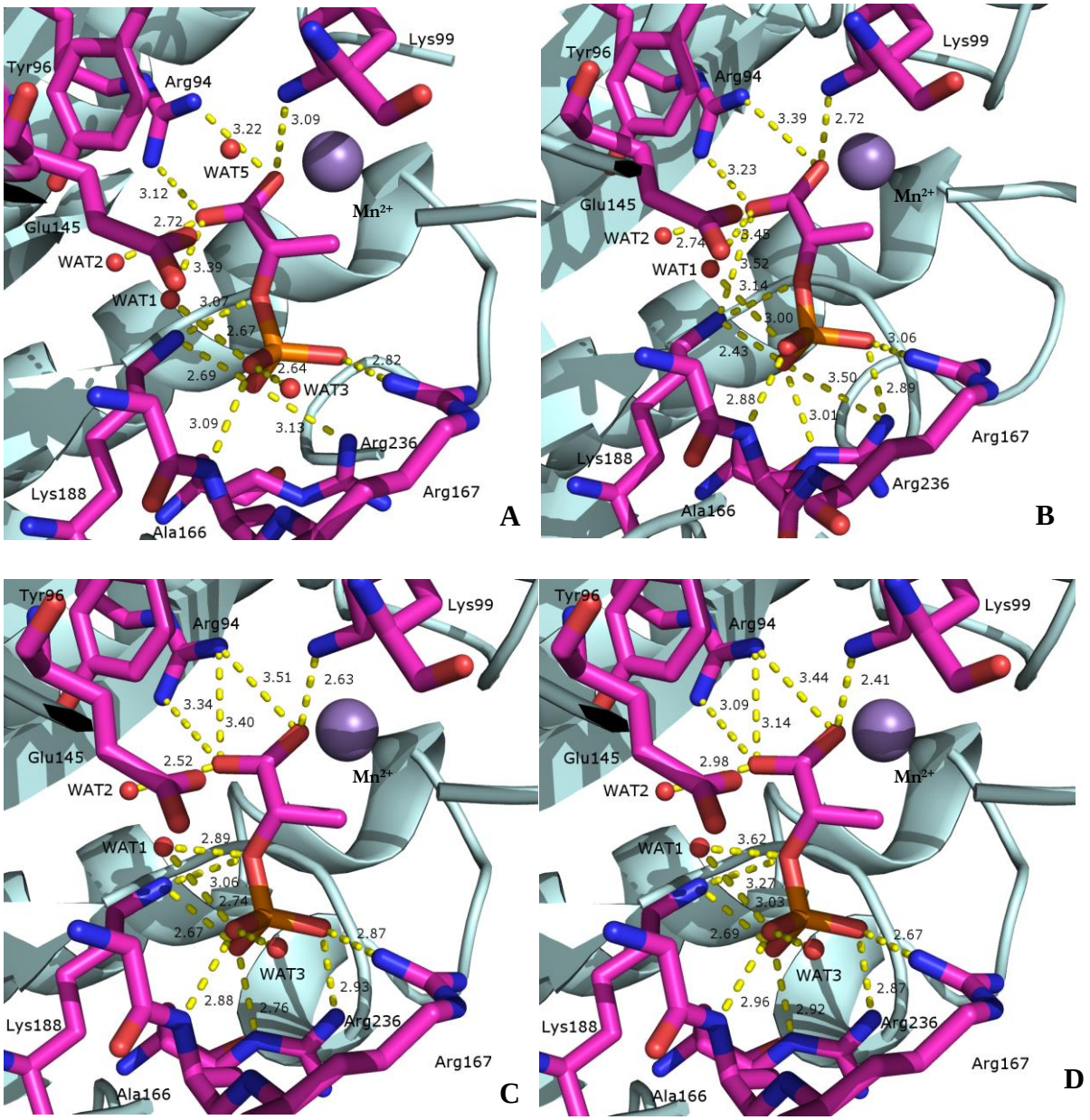


Figure S8. Interatomic distances of (S)-phospholactate 8 bound at the active site of NmeDAH7PS. The four chains of NmeDAH7PS are indicated: top left: chain A, top right: chain B, bottom left: chain C and bottom right: chain D. Mn^{2+} is shown as a purple sphere in all panels.

7.3.1 Inhibition Kinetics of three aromatic amino acids

Standard assays were performed, which consisted of E4P (250 µM), PEP (100 µM), $MnSO_4$ (100 µM) and the amino acid; L-Phe, L-Tyr or L-Trp (50 µM – 1 mM) in 50 mM BTP, pH 6.8. Assays were carried out at 25 °C and were initiated by the addition of *Nme*DAH7PS. All Assays were performed in triplicate and averaged.

7.3.2. Isothermal titration calorimetry

Binding of Phe to *Nme*DAH7PS at pH 7 was measured by ITC using a VP-ITC unit operating at 298 K (Microcal, GE Healthcare). Before use, the protein was buffer exchanged against binding buffer [0.5 mM $MnSO_4$ in 50 mM BTP buffer (pH 7)] and all solutions were degassed in a vacuum. Protein concentration was measured by UV absorption immediately before titrations were started. The titrations were comprised of 28 injections, one 2 µL injection followed by 27 x 10 µL injections of Phe. The initial datum point was routinely deleted to allow for diffusion of ligand across the needle tip during the equilibration period. A heat of dilution experiment was measured independently and subtracted from the integrated data before curve fitting in Origin 7.0. For the binding of Phe to *Nme*DAH7PS, 16 µM of *Nme*DAH7PS was used and the syringe contained 1.2 mM Phe. The data were fitted with the two-site sequential-binding model supplied by Micro-Cal.

Binding of Trp to *Nme*DAH7PS was completed using an ITC200 unit operating at 298 K (MicroCal; GE Healthcare) using the same sample preparation as described above however samples were not degassed in this case. The titrations comprised of one 0.5 µL injection followed by 19 injections of 2 µL. The initial datum point was routinely deleted to allow for diffusion of ligand across the needle tip during the equilibration period. For the binding of Trp to *Nme*DAH7PS, 30 µM of *Nme*DAH7PS was used and the syringe contained 20 mM Trp. The data could not be fitted using the models supplied by Micro-Cal.

7.3.3 Crystal soaking, co-crystallisation, structure determination and refinement

Soaking of the aromatic amino acids into crystals of *Nme*DAH7PS was achieved using the established procedures outlined in (section 7.1.32). Tyr was difficult to dissolve at high concentrations and therefore required the addition of a small volume of base (sodium hydroxide) to the crystallisation condition before soaking. Co-crystallisation of ligands at both active and allosteric sites occurred making up the reported crystallisation condition with 1.5mM Phe and 1 mM PEP. Several hanging drops were prepared using the previously established method (*Nme*DAH7PS ~10 mg/mL) per well and an already formed *Nme*DAH7PS crystal in the absence of these ligand concentrations was used to seed into these drops. A seeding dilution series was carried out across all drops until an appropriate crystal had formed using procedures outlined in (section 7.1.31).

The X-ray diffraction data for the liganded structures were collected at the Australian Synchrotron on the Macro (MX1) and Micro Crystallography (MX2) beamlines. All four liganded structures (Phe, Trp, Tyr, Phe+PEP) crystallised like *Nme*DAH7PS^{WT} in a monoclinic space group $P\ 12_11$ and diffracted to 2.19 Å, 2.23 Å, 2.31 Å, and 2.05 Å, respectively. The results are summarised in Table 3.2.3.1, along with key structure refinement details.

All structures were phased by molecular replacement using Phaser MR[116], with the same R_{free} reflection set.

EXPERIMENTAL FOR CHAPTER 4 PART I

The research described in chapter 4 part I was the subject of a publication before submission of this thesis. The supplementary material for this publication reporting experimental details appended with the permission of the copyright owners and can be accessed at:

SUPPLEMENTAL DATA

The functional unit of *Neisseria meningitidis* 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase is dimeric

Logan C Heyes[1]*, Penelope J Cross[1]*, Shiwen Zhang, Ali Reza Nazmi[1] and Emily J Parker[1]

[1]Biomolecular Interaction Centre and Department of Chemistry, University of Canterbury, Christchurch, New Zealand

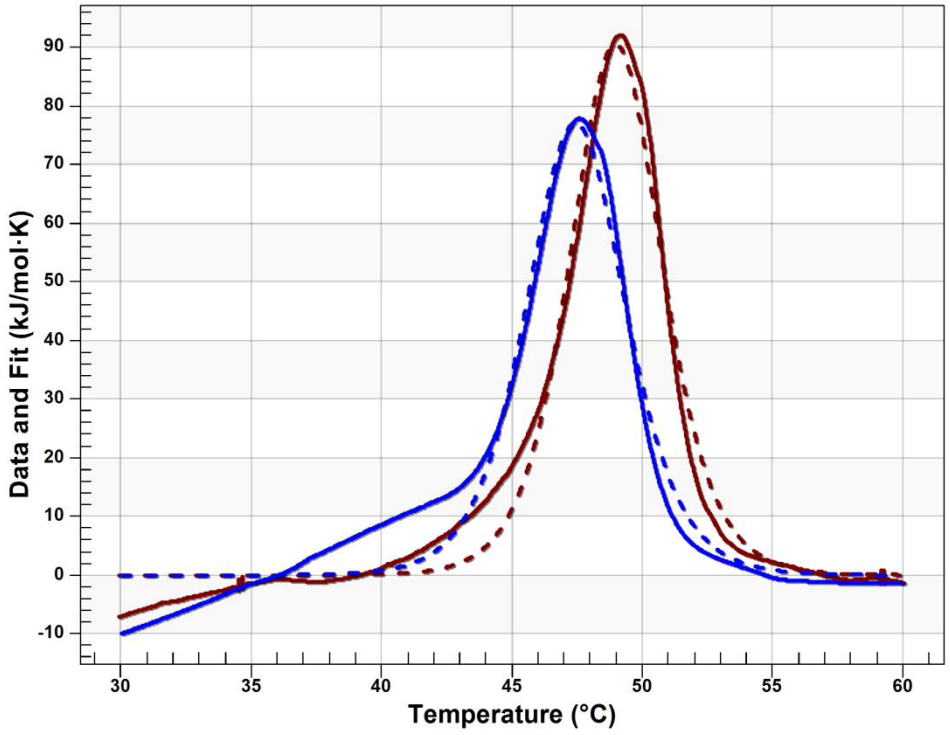


Figure S1. Differential scanning calorimetry showing thermal stability of *Nme*DAH7PS^{WT} and *Nme*DAH7PS^{R126S}. *Nme*DAH7PS^{WT} (brown) and *Nme*DAH7PS^{R126S} (blue). Dashed lines indicate the model that the data has been fit to.

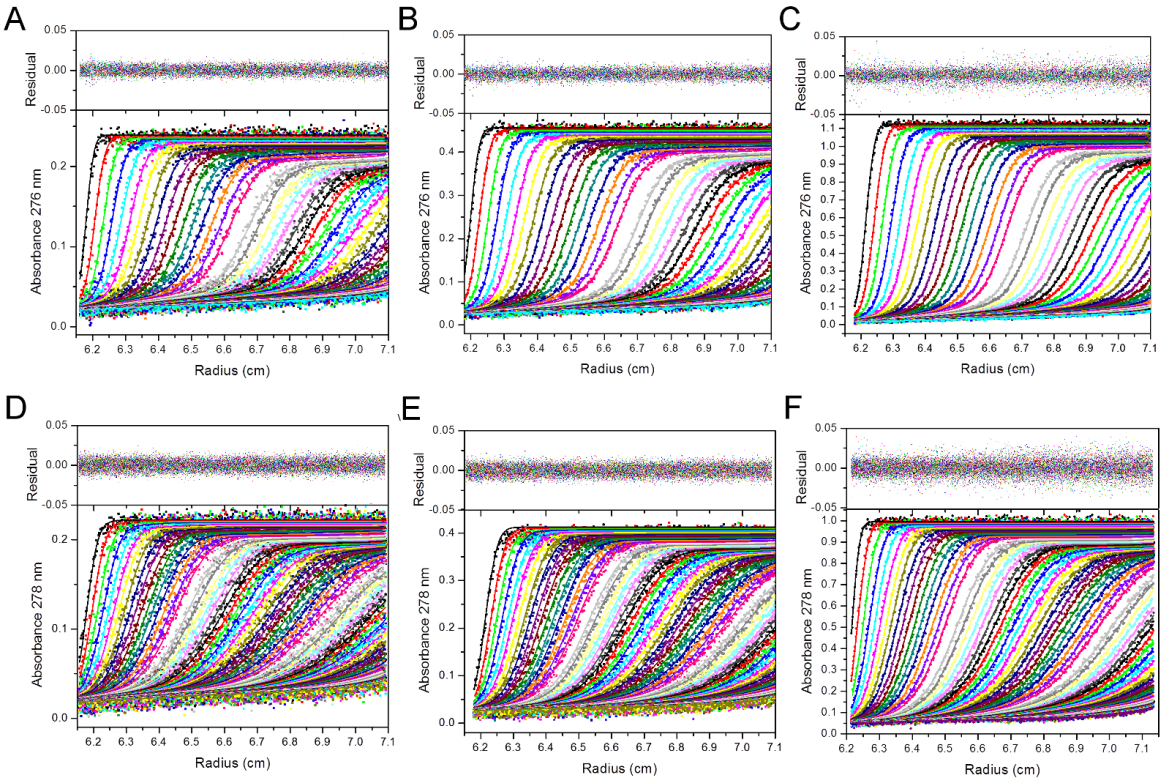


Figure S2. Sedimentation velocity data collected at 20 ° C and 50,000 rpm. (A), (B) and (C) are the data from NmeDAH7PS^{WT} collected at 0.2, 0.4 and 1.2 mg.mL^{-1} and (D), (E) and (F) are the data from NmeDAH7PS^{R126S} collected at 0.2, 0.4 and 1.2 mg.mL^{-1}. The top panel in each section shows the sedimentation velocity data, the size-distribution best fit and residuals for the data fits.

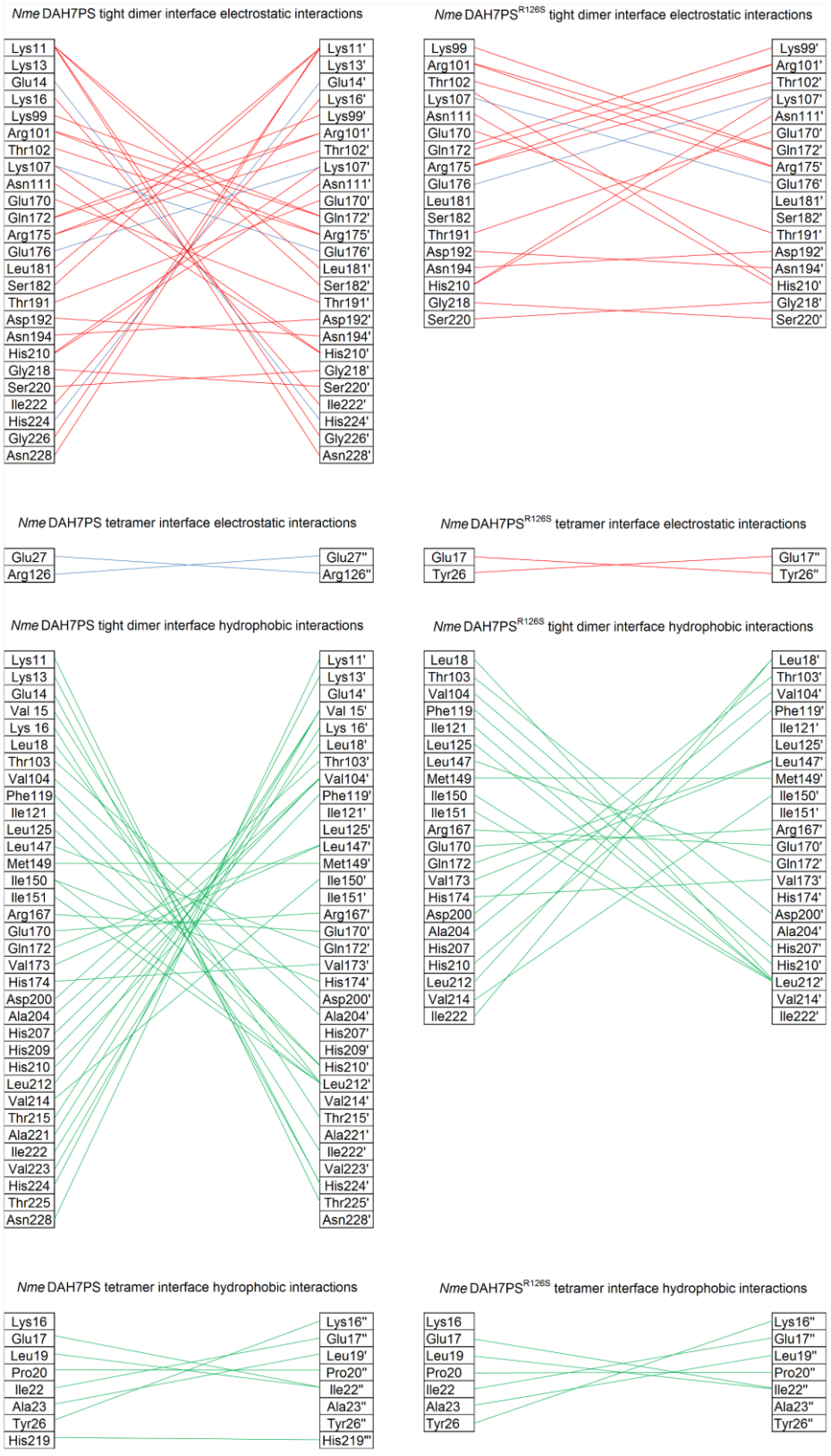


Figure S3. Interface analysis alignment of, left, NmeDAH7PS^{WT} and right, NmeDAH7PS^{R126S}. Green lines indicate hydrophobic interactions at the interface, red lines indicate hydrogen bonds at the interface and blue lines indicate salt-bridges at the interface.

EXPERIMENTAL FOR CHAPTER 4 PART II

7.4.1 Site-directed mutagenesis

The constructs for the variant protein in this chapter were created by site-directed mutagenesis using *Nme*DAH7PS^{WT} construct plasmids as the templates.

*Nme*DAH7PS H219A

For: 5′ TGG ACA ATG GCG GAA GCC CCG GCC TTG GTT AC

Rev: 5′ GTA ACC AAG GCC GGG GCT TCC GCC ATT GTC CA

The PCR products were transformed into Top10 cells. Plasmids of the expected size, as determined by agarose gel, were DNA sequenced to identify those containing the successful mutation and transformed into BL21 Star^{TM} (DE3) cells.
Expression and purification were carried out as previously described for *Nme*DAH7PS^{WT}.

7.4.2 Kinetics and Inhibition

Standard assays to determine kinetic parameters were carried out according to methods described in 7.1.27. The assays for determining the kinetic parameters kept the concentration of one substrate constant while varying that of the other and vice versa. Reactions were initiated with 2 µl of 1mg/mL *Nme*DAH7PS^{His219Ala}. To determine the $K_{M\ (PEP)}$ the E4P concentration was kept at 300 µM. The concentration of PEP was varied from 2.5-160 µM. To determine $K_{M\ (E4P)}$, the PEP concentration was kept at 100 µM and the E4P concentration was varied from 5-180 µM.

Inhibition assays were performed, which consisted of E4P (500 µM), PEP (100 µM), $MnSO_4$ (100 µM) and the amino acid Phe (50 µM – 1 mM) in 50 mM BTP, pH 6.8. Assays were carried out at 25 °C and were initiated by the addition of *Nme*DAH7PS. All assays were performed in triplicate and averaged.

7.4.3 Isothermal titration calorimetry

ITC was carried out using procedures described in (section 7.1.29). For *Nme*DAH7PS^{His219Ala} the protein concentration in the cell was 20 µM and the Phe concentration in the syringe was 1.2 mM. A heat of dilution experiment was measured independently and subtracted from the integrated data before curve fitting in Origin 7.0. The data were fitted with the two-site sequential-binding model supplied by Micro-Cal.

7.4.4 Analytical size exclusion chromatography

To determine the quaternary structure of the *Nme*DAH7PS^{His219Ala} variant, it was applied to a WYATT© size exclusion chromatography column for multi-angle light scattering. Comparison was made to *Nme*DAH7PS^{WT} as a void volume could not be obtained. 500 µl of 0.7 mg/mL was applied to the column.

7.4.5 Crystallisation, structure determination and refinement

Micro-seeding of *Nme*DAH7PS^{His219Ala} occurred using procedures outlined in (section 7.1.31). Soaking of the aromatic amino acids into crystals of *Nme*DAH7PS was achieved using the established procedures outlined in (section 7.1.32). The X-ray diffraction data for both the liganded and apo structures were collected at the Australian Synchrotron on the Macro (MX1) Crystallography beamlines. The liganded and apo structures crystallised like *Nme*DAH7PS^{WT} protein in a monoclinic space group $P\ 12_11$ and diffracted to 2.00 Å and 2.21 Å, respectively. The results are summarised in Table 4.3.5.1, along with key structure refinement details.

All structures were phased by molecular replacement using Phaser MR[116], with the same R_{free} reflection set.

7.5.1 Site-directed mutagenesis

The constructs for the variant proteins in this chapter were created by site-directed mutagenesis using *Nme*DAH7PS^{WT} construct plasmids as the templates.

*Nme*DAH7PS E98A

For: 5′ TCG TCC TCG GCT TCG CGA AAT AAA CGC GCA TCA
Rev: 5′ TGA TGC GCG TTT ATT TCG CGA AGC CGA GGA CGA

*Nme*DAH7PS K107A

For: 5′ GGT CGT TAA TCA AAC CTG CCC AAC CCA CCG TCG TCC
Rev: 5′ GGA CGA CGG TGG GTT GGG CAG GTT TGA TTA ACG ACC

*Nme*DAH7PS Q172A

For: 5′ GCG GAC GAC CGA AAG CGC AGT TCA CCG CCA ATT G
Rev: 5′ CAA TTC GCG GTG AAC TGC GCT TTC GGT CGT CCG C

*Nme*DAH7PS E176A

For: 5′ GCC AAG TTC ACC GCG CGT TGG CAA GCG GGC TG
Rev: 5′ CAG CCC GCT TGC CAA CGC GCG GTG AAC TTG GC

*Nme*DAH7PS E176Q

For: 5′ AGC CAA GTT CAC CGC CAG TTG GCA AGC GGG CTG
Rev: 5′ CAG CCC GCT TGC CAA CTG GCG TGT AAC TTG GCT

*Nme*DAH7PS F211A

For: 5′ CCT TGG TTA CAG ACA GGG CAT GAT GCG AAT GGC TCG CC
Rev: 5′ GGC GAG CCA TTC GCA TCA TGC CCT GTC TGT AAC CAA GG

*Nme*DAH7PS V223A

For: 5′ GGC ATT CCG CCA TTG CCC ATA CCG GCG GCA A
Rev: 5′ TTG CCG CCG GTA TGG GCA ATG GCG GAA TGC C

The PCR products were transformed into Top10 cells. Plasmids of the expected size, as determined by agarose gel, were DNA sequenced to identify those containing the successful mutation and transformed into BL21 Star^{TM} (DE3) cells.

Expression and purification were carried out as previously described for *Nme*DAH7PS^{WT}.

7.5.2 Kinetics and Inhibition

Standard assays to determine kinetic parameters were carried out according to methods described in 7.1.27. The assays for determining the kinetic parameters kept the concentration of one substrate constant while varying that of the other and vice versa. Reactions involving *Nme*DAH7PS^{Phe211Ala}, *Nme*DAH7PS^{Glu98Ala} and *Nme*DAH7PS^{Glu176Ala} were initiated with 2 µl of 1 mg/mL. Reactions involving *Nme*DAH7PS^{Gln172Ala}, *Nme*DAH7PS^{Glu176Gln} and *Nme*DAH7PS^{Val223Ala} were initiated with 2 µl of 0.6 mg/mL.

To determine the $K_{M\ (PEP)}$ the E4P concentration was kept at 300 µM for all variants except *Nme*DAH7PS^{Gln172Ala}, *Nme*DAH7PS^{Glu176Gln} and *Nme*DAH7PS^{Phe211Ala} for which the E4P concentration was 500 µM. The concentration of PEP was varied from 1-80 µM for *Nme*DAH7PS^{Glu98Ala}, 4-128 µM for *Nme*DAH7PS^{Gln172Ala} and *Nme*DAH7PS^{Glu176Gln}, 4-64 µM for *Nme*DAH7PS^{Glu176Ala}, 2.5-80 µM for *Nme*DAH7PS^{Phe211Ala} and 5-120 µM for *Nme*DAH7PS^{Val223Ala}. To determine $K_{M\ (E4P)}$ for these mutants, the PEP concentration was kept at 100 µM and the E4P concentration was varied from 10-240 µM for *Nme*DAH7PS^{Glu98Ala}, 10-160 µM for *Nme*DAH7PS^{Glu176Ala} and *Nme*DAH7PS^{Val223Ala}, 6.25-250 µM for *Nme*DAH7PS^{Phe211Ala}, 12-384 µM for *Nme*DAH7PS^{Glu176Gln}, and 10-320 µM for *Nme*DAH7PS^{Gln172Ala}.

Inhibition assays were performed, which consisted of E4P (500 µM), PEP (100 µM), $MnSO_4$ (100 µM) and the amino acid Phe (50 µM – 1 mM) in 50 mM BTP, pH 6.8. Assays were carried out at 25 °C and were initiated by the addition of *Nme*DAH7PS. All assays were performed in triplicate and averaged.

7.5.3 Isothermal titration calorimetry

ITC was carried out using procedures described in (section 7.1.29). for *Nme*DAH7PS^{Gln172Ala}, *Nme*DAH7PS^{Glu176Gln} and *Nme*DAH7PS^{Phe211Ala} the protein concentration in the cell was 16 µM

and the Phe concentration in the syringe was 1.2 mM. For *Nme*DAH7PS^{Glu176Ala} the protein concentration in the cell was 22 µM and the Phe concentration in the syringe was 0.2 mM. For *Nme*DAH7PS^{Lys107Ala} the protein concentration in the cell was 21 µM and the Phe concentration in the syringe was 0.1 mM. For *Nme*DAH7PS^{Glu98Ala} the protein concentration in the cell was 18 µM and the Phe concentration in the syringe was 1.2 mM. For *Nme*DAH7PS^{Val223Ala} the protein concentration in the cell was 20 µM and the Phe concentration in the syringe was 20 mM. A heat of dilution experiment was measured independently and subtracted from the integrated data before curve fitting in Origin 7.0. The data were fitted with a one-site or two-site sequential-binding model supplied by MicroCal.

7.5.4 Analytical size exclusion chromatography

To determine the quaternary structure of *Nme*DAH7PS variants, they were all applied to a WYATT© size exclusion chromatography column for multi-angle light scattering. Comparison was made to *Nme*DAH7PS^{WT} as a void volume could not be obtained. 500 µl of 0.6-0.7 mg/mL of variant proteins was applied to the column.

7.5.5 Crystallisation, structure determination and refinement

Micro-seeding of all *Nme*DAH7PS variant crystals occurred using procedures outlined in (section 7.1.31). Soaking of Phe or PEP into crystals of *Nme*DAH7PS was achieved using the established procedures outlined in (section 7.1.32). The X-ray diffraction data for both the liganded, PEP and apo structures were collected at the Australian Synchrotron on the Macro (MX1) and micro (MX2) Crystallography beamlines. The Phe-bound *Nme*DAH7PS^{Glu176Ala} structure crystallised like the *Nme*DAH7PS^{WT} in a monoclinic space group *P* 12_11, While the apo structure crystallised in an orthorhombic space group *P* $2_12_12_1$. The two structures were refined to 2.04 Å and 2.40 Å, respectively. The Phe+PEP-bound and PEP-bound *Nme*DAH7PS^{Glu98Ala} structures crystallised like the *Nme*DAH7PS^{WT} in a monoclinic space group *P* 12_11 and were refined to 2.42 Å and 2.50 Å, respectively. The Phe-bound and PEP-bound *Nme*DAH7PS^{Glu176Gln} structures crystallised like the *Nme*DAH7PS^{WT} in a monoclinic space group *P* 12_11 and were refined to 2.70 Å and 1.87 Å, respectively. The Phe-bound and PEP-bound *Nme*DAH7PS^{Gln172Ala} structures crystallised like the *Nme*DAH7PS^{WT} in a monoclinic space group *P* 12_11 and were refined to 2.30 Å and 2.68 Å, respectively. The Phe+PEP-bound

*Nme*DAH7PS^{Lys107Ala} structure crystallised like the *Nme*DAH7PS^{WT} in a monoclinic space group *P* 12_11 and was refined to 1.75 Å. All results are summarised in table 5.3.4.1 and 5.3.4.2.

The *Nme*DAH7PS^{Phe211Ala} structure crystallised in an orthorhombic space group *P* $2_12_12_1$ and was refined to 2.35 Å. The Phe-bound *Nme*DAH7PS^{Val223Ala} structure crystallised in an orthorhombic space group *P* 2_122_1 and only contained the dimer in the unit cell, unlike any of the *Nme*DAH7PS structures collected to date. The PEP-bound *Nme*DAH7PS^{Val223Ala} structure crystallised like the *Nme*DAH7PS^{WT} in a monoclinic space group *P* 12_11. The structures were refined to a resolution of 1.70 Å and 1.84 Å respectively. All results are summarised in table 5.3.3.1.

All structures were phased by molecular replacement using Phaser MR[116], with the same R_{free} reflection set.

REFERENCES

1. Shulman, S. T. The History of Pediatric Infectious Diseases. *Pediatr. Res.* **55,** 163–176 (2004).

2. Peltola, H. Meningococcal disease: still with us. *Rev. Infect. Dis.* **5,** 71–91 (1983).

3. Schuchat, a, Robinson, K., Wenger, J. D., Harrison, L. H., Farley, M., Reingold, a L., Lefkowitz, L. & Perkins, B. a. Bacterial meningitis in the United States in 1995. Active Surveillance Team. *N. Engl. J. Med.* **337,** 970–976 (1997).

4. Brouwer, M. C., Tunkel, a. R. & van de Beek, D. Epidemiology, Diagnosis, and Antimicrobial Treatment of Acute Bacterial Meningitis. *Clin. Microbiol. Rev.* **23,** 467–492 (2010).

5. Teyssou, R. & Muros-Le Rouzic, E. Meningitis epidemics in Africa: A brief overview. *Vaccine* **25,** 7–11 (2007).

6. Nikulin, J., Panzner, U., Frosch, M. & Schubert-Unkmeir, A. Intracellular survival and replication of Neisseria meningitidis in human brain microvascular endothelial cells. *Int. J. Med. Microbiol.* **296,** 553–558 (2006).

7. Perkins, B., Rosenstein, N., Stephens, D., Popovic, T. & Hughes, J. Meningicoccal disease. *N. Engl. J. Med.* **344,** 1378–1388 (2001).

8. World health organisation. Meningococcal meningitis fact sheet No. 141. (2015). at <Http://www.who.int/mediacentre/factsheets/fs141/en/>

9. Tzeng, Y.-L., Martin, L. E. & Stephens, D. S. Environmental survival of *Neisseria meningitidis*. *Epidemiol. Infect.* **142,** 187–90 (2014).

10. Rouphael, N. G. & Stephens, D. S. *Neisseria meningitidis*: Biology, microbiology, and epidemiology. *Methods Mol. Biol.* **799,** 1–20 (12AD).

11. Merz, A. J. & So, M. Interactions of pathogenic neisseriae with epithelial cell membranes. *Annu. Rev. Cell Dev. Biol.* **16,** 423–457 (2000).

12. Nassif, X. Interaction mechanisms of encapsulated meningococci with eucaryotic cells: what does this tell us about the crossing of the blood–brain barrier by *Neisseria meningitidis*? *Curr. Opin. Microbiol.* **2,** 71–77 (1999).

13. Vogel, U. & Frosch, M. Mechanisms of neisserial serum resistance. *Mol. Microbiol.* **32,** 1133–1139 (1999).

14. Esposito, S., Castellazzi, L., Bosco, A., Musio, A. & Stoddard, J. Use of a multicomponent, recombinant, meningococcal serogroup B vaccine (4CMenB) for bacterial meningitis prevention. *Immunotherapy* **6,** 395–408 (2014).

15. Jafri, R. Z., Ali, A., Messonnier, N. E., Tevi-Benissan, C., Durrheim, D., Eskola, J., Fermon, F., Klugman, K. P., Ramsay, M., Sow, S., Zhujun, S., Bhutta, Z. a & Abramson, J. Global epidemiology of invasive meningococcal disease. *Popul. Health Metr.* **11,** 17 (2013).

16. Granoff, D. M. Review of Meningococcal Group B Vaccines. *Clin. Infect. Dis.* **50,** S54–S65 (2010).

17. Chaudhuri, A., Martin, P. M., Kennedy, P. G. E., Andrew Seaton, R., Portegies, P., Bojar, M. & Steiner, I. EFNS guideline on the management of community-acquired bacterial meningitis: report of an EFNS Task Force on acute bacterial meningitis in older children and adults. *Eur. J. Neurol.* **15,** 649–659 (2008).

18. Bentley, R. The Shikimate Pathway - A Metabolic Tree with Many Branches. *Crit. Rev. Biochem. Mol. Biol.* **25,** 307–384 (1990).

19. Roberts, F., Roberts, C., Johnson, J., Kyle, D., Krell, T., Coggins, J., Coombs, G., Milhous, W., Tzipori, S., Ferguson, D., Chakrabarti, D. & McLeod, R. Evidence for the shikimate pathway in apicomplexan parasites. *Nature* **393,** 801–805 (1998).

20. Light, S. H. & Anderson, W. F. The diversity of allosteric controls at the gateway to aromatic amino acid biosynthesis. *Protein Sci.* **22,** 395–404 (2013).

21. Light, S. H., Halavaty, A. S., Minasov, G., Shuvalova, L. & Anderson, W. F. Structural analysis of a 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase with an N-terminal chorismate mutase-like regulatory domain. *Protein Sci.* **21,** 887–95 (2012).

22. Shumilin, I., Bauerle, R. & Kretsinger, R. Crystal structure of phenylalanine-regulated 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Escherichia coli*. *Structure* **7,** 865–875 (1999).

23. J Sikorski, and K. G. Understanding Glyphosate's Molecular Mode of Action with EPSP Synthase: Evidence Favoring an Allosteric Inhibitor Mode. *Acc. Chem. Res.* **30,** 2–8 (1997).

24. Blackmore, N. J., Reichau, S., Jiao, W., Hutton, R. D., Baker, E. N., Jameson, G. B. & Parker, E. J. Three Sites and You Are Out: Ternary Synergistic Allostery Controls Aromatic Amino Acid Biosynthesis in *Mycobacterium tuberculosis*. *J. Mol. Biol.* **425,** 1582–1592 (2013).

25. Cross, P. J., Dobson, R. C. J., Patchett, M. L. & Parker, E. J. Tyrosine latching of a regulatory gate affords allosteric control of aromatic amino acid biosynthesis. *J. Biol. Chem.* **286,** 10216–10224 (2011).

26. Cross, P. J., Pietersma, A. L., Allison, T. M., Wilson-Coutts, S. M., Cochrane, F. C. & Parker, E. J. *Neisseria meningitidis* expresses a single 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase that is inhibited primarily by phenylalanine. *Protein Sci.* **22,** 1087–1099 (2013).

27. Jiao, W., Hutton, R. D., Cross, P. J., Jameson, G. B. & Parker, E. J. Dynamic cross-talk among remote binding sites: the molecular basis for unusual synergistic allostery. *J. Mol. Biol.* **415,** 716–26 (2012).

28. Shumilin, I. A., Zhao, C., Bauerle, R. & Kretsinger, R. H. Allosteric Inhibition of 3-Deoxy-D-arabino-heptulosonate 7-phosphate Synthase Alters the Coordination of Both Substrates. *J. Mol. Biol.* **320,** 1147–1156 (2002).

29. Blackmore, N. J., Nazmi, A. R., Hutton, R. D., Webby, M. N., Baker, E. N., Jameson, G. B. & Parker, E. J. Complex Formation between Two Biosynthetic Enzymes Modifies the Allosteric Regulatory Properties of Both. *J. Biol. Chem.* **290,** 18187–18198 (2015).

30. Parker, E. J., Bulloch, E. M. M., Jameson, G. B. & Abell, C. Substrate Deactivation of Phenylalanine-Sensitive 3-Deoxy- D-*arabino*-heptulosonate 7-phosphate synthase. *Biochemistry* **40,** 14821–14828 (2001).

31. Shumulin, I. A., Bauerle, R. & Kretsinger, R. H. The High-Resolution Structure of 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase Reveals a Twist in the Plane of Bound Phosphoenolpyruvate. *Biochemistry* **42,** 3766–3776 (2003).

32. Williamson, R., Pietersma, A., Jameson, G. & Parker, E. Stereospecific deuteration of 2-deoxyerythrose 4-phosphate using 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *J. Bioorganic Med. Chem. Lett.* **15,** 2339–2342 (2005).

33. V Konig, Braus, G., Schneider, T. & Pfeil, A. Substrate and metal complexes of 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Saccharomyces cerevisiae* provide new insights into the catalytic mechanism. *J. Mol. Biol.* **337,** 675–690 (2004).

34. Takahashi, M. & Chan, W. W. Separation and properties of isozymes of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase from *Saccharomyces cerevisiae*. *Can. J. Biochem.* **49,** 1015–1025 (1971).

35. Webby, C., Baker, H., Lott, S., Baker, E. & Parker, E. The Structure of 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase from *Mycobacterium tuberculosis* Reveals a Common Catalytic Scaffold and Ancestry for Type I and Type II Enzymes. *J. Mol. Biol.* **354,** 927–939 (2005).

36. Schofield, L., Patchett, M. & Parker, E. Expression, purification, and characterization of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*. *Protein Expr. Purif.* **34,** 17–27 (2004).

37. Wu, J., Howe, D. L. & Woodard, R. W. *Thermotoga maritima* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase: the ancestral eubacterial DAHP synthase? *J. Biol. Chem.* **278,** 27525–31 (2003).

38. Ahn, M., Pietersma, A. L., Schofield, L. R., Parker, E. J., M Ahn L Schofield and E Parker, A. P., Ahn, M., Pietersma, A. L., Schofield, L. R. & Parker, E. J. Mechanistic divergence of two closely related aldol-like enzyme-catalysed reactions. *Org. Biomol. Chem.* **3,** 4046–9 (2005).

39. Allison, T. M., Hutton, R. D., Jiao, W., Gloyne, B. J., Nimmo, E. B., Jameson, G. B. & Parker, E. J. An Extended β7α7 Substrate-Binding Loop Is Essential for Efficient Catalysis by 3-Deoxy-D-*manno*-Octulosonate 8-Phosphate Synthase. *Biochemistry* **50,** 9318–9327 (2011).

40. Schofield, L. R., Andersen, B. F., Patchett, M. L., Norris, G. E., Jameson, G. B. & Parker, E. J. Substrate Ambiguity and Crystal Structure of *Pyrococcus furiosus* 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase: An Ancestral 3-Deoxyald-2-ulosonate-phosphate Synthase? *Biochemistry* **44,** 11950–11962 (2005).

41. Kona, F., Xu, X., Martin, P., Kuzmic, P. & Gatti, D. L. Structural and Mechanistic Changes along an Engineered Path from Metallo to Nonmetallo 3-Deoxy-D-*manno*-octulosonate 8-Phosphate Synthases. *Biochemistry* **46,** 4532–4544 (2007).

42. Allison, T. M., Hutton, R. D., Cochrane, F. C., Yeoman, A., Jameson, B. & Parker, E. J. Targeting the Role of a Key Conserved Motif for Substrate Selection. *Biochemistry* **50,** 3686–3695 (2011).

43. Webby, C. J. Structural & functional characterisation of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Helicobacter pylori* & *Mycobacterium tuberculosis*. *Dr. Philos. thesis* **Massey Uni,** (2006).

44. Gosset, G., Bonner, C. A. & Jensen, R. Microbial origin of plant-type 2-keto-3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and tryptophan-regulated enzyme from *Xanthomonas campestris*. *J. Bacteriol.* **183,** 4061–4070 (2001).

45. Silakowski, B., Kunze, B. & Muller, R. Stigmatella aurantiaca Sg a15 carries genes encoding type I and type II 3-deoxy- D -*arabino*-heptulosonate 7-phosphate synthases : involvement of a type II synthase in aurachin biosynthesis. *Arch. Microbiol.* **173,** 403–411 (2000).

46. Kloosterman, H., Hessels, G. I., Vrijbloed, J. W., Euverink, G. J. & Dijkhuizen, L. (De) regulation of key enzyme steps in the shikimate pathway and phenylalanine-specific pathway of the actinomycete *Amycolatopsis methanolica*. *Microbiiology* **149,** 3321–3330 (2003).

47. Nazmi, A. R., Othman, M., Lang, E. J. M., Bai, Y., Allison, T. M., Panjokar, S., Arcus, V. L. & Parker, E. J. Tipping an allosteric hat: The catalytic regulatory domain caps the active site of 3-deoxy-D-*arabino* heptulosonate 7-phosphate-synthase from *Geobacillus sp*. *Unpublished* (2015).

48. Webby, C. J., Jiao, W., Hutton, R. D., Blackmore, N. J., Baker, H. M., Baker, E. N., Jameson, G. B. & Parker, E. J. Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid biosynthesis in Mycobacterium tuberculosis. *J. Biol. Chem.* **285,** 30567–30576 (2010).

49. Tran, D., Pietersma, A., Schofield, L., Rost, M., Jameson, G. & Parker, E. Investigating the role of the hydroxyl groups of substrate erythrose 4-phosphate in the reaction catalysed by the first enzyme of the shikimate pathway. *Bioorganic Med. Chem. Lett.* **21,** 6838–6841 (2011).

50. Wagner, T., Shumulin, I., Bauerle, R. & Kretsinger, R. Structure of 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase from *Escherichia coli*: Comparison of the Mn2+ 2-Phosphoglycolate and the Pb2+ 2-Phosphoenolpyruvate Complexes and Implications for Catalysis. *J. Mol. Biol.* **301,** 389–399 (2000).

51. Walker, S., Cumming, H. & Parker, E. Substrate and reaction intermediate mimics as inhibitors of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *J. Org. Biomol. Chem.* **7,** 3031–3035 (2009).

52. Grison, C., Petek, S., Finance, C. & Coutrot, P. Synthesis and antibacterial activity of mechanism-based inhibitors of KDO8P synthase and DAH7P synthase. *Carbohydr. Res.* **340,** 529–37 (2005).

53. Walker, S. & Parker, E. Synthesis and Evaluation of a mechanism based inhibitor of a 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *J. Bioorganic Med. Chem. Lett.* **16,** 2951–2954 (2006).

54. Reichau, S., Jiao, W., Walker, S., Hutton, R., Baker, E. & Parker, E. Potent Inhibitors of a Shikimate Pathway Enzyme from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **286,** 16197–161207 (2011).

55. Deleo, A. B., Dayan, J. & Sprinson, D. B. Purification and kinetics of tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthetase from *Salmonella*. *J. Biol. Chem.* **248,** 2344–2353 (1973).

56. Simpson, R. I. & Davidson, B. E. Studies on 3 deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase(phe) from *Escherichia coli* K12. *Eur. J. Biochem.* **70,** 493–500 (1976).

57. Schoner, R. & Herrmann, K. M. 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase-Purification, properties and kinetics of the tyrosine-sensitive isoenzyme from *Eschericia coli*. *J. Biol. Chem.* **251,** 5440–5447 (1976).

58. Deleo, A. B. & Sprinson, D. B. Mechanism of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthetase. *Biochem. Biophys. Res. Commun.* **32,** 873–877 (1968).

59. Floss, H. G., Onderka, D. K. & Carroll, M. Stereochemistry of the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase reaction and the chorismate synthetase reaction. *J. Biol. Chem.* **247,** 736–744 (1972).

60. Onderka, D. & Floss, H. G. Steric course of the Chorismate Synthetase Reaction and the 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) Synthetase Reaction. *J. Amercian Chem. Soc.* **81,** 5894–5896 (1969).

61. Onderka, D. K. & Floss, H. G. Stereospecificity of the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase reaction. *Biochem. Biophys. Res. Commun.* **35,** 801–804 (1969).

62. Zhou, L., Wu, J., Janakiraman, V., Shumulin, I. A., Bauerle, R., Kretsinger, R. H. & Woodard, R. W. Structure and characterization of the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Aeropyrum pernix*. *Bioorg. Chem.* **40,** 79–86 (2012).

63. Shumulin, I. A., Bauerle, R., Wu, J., Woodard, R. W. & Kretsinger, R. H. Crystal Structure of the Reaction Complex of 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase from *Thermotoga maritima* Refines the Catalytic Mechanism and Indicates a New Mechanism of Allosteric Regulation. *J. Mol. Biol.* **341,** 455–466 (2004).

64. Reichau, S. & Parker, E. J. Active site plasticity of a critical enzyme from *Mycobacterium tuberculosis*. *R. Soc. Chem.* **3,** 3209–3213 (2013).

65. Kona, F., Tao, P., Martin, P., Xu, X. & Gatti, D. Electronic Structure of the Metal Center in the Cd2+, Zn2+, and Cu2+ Substituted Forms of KDO8P Synthase: Implications for Catalysis. *Biochemistry* **48,** 3610–3630 (2009).

66. Tao, P., Gatti, D. & Schlegel, H. The Energy Landscape of 3-Deoxy-D-*manno*-octulosonate 8-Phosphate Synthase. *Biochemistry* **48,** 11706–11714 (2009).

67. Walker, S. R. Design, synthesis and characterisation of inhibitors of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. *PhD Thesis* **University,** (2007).

68. Schwartz, S. & Schramm, V. Enzymatic transition states and dynamic motion in barrier crossing. *Nat. Chem. Biol.* **5,** 551–558 (2009).

69. Schramm, V. L. Enzymatic transition states and transition state analogues. *Curr. Opin. Struct. Biol.* **15,** 604–13 (2005).

70. Schramm, V. L. Transition States, analogues, and drug development. *ACS Chem. Biol.* **8,** 71–81 (2013).

71. Onderka, D. & Floss, H. Steric Course of the Chorismate Synthetase Reaction and the 3-Deoxy-D-*arabino*-heptulosonate 7-Phosphate (DAHP) Synthetase Reaction. *J. Am. Chem. Soc.* **91,** 5894–5896 (1969).

72. Cross, P. J., Allison, T. M., Dobson, R. C. J., Jameson, G. B. & Parker, E. J. Engineering allosteric control to an unregulated enzyme by transfer of a regulatory domain. *PNAS* **110,** 2111–2116 (2013).

73. Webby, C. J., Lott, J. S., Baker, H. M., Baker, E. N. & Parker, E. J. Crystallization and preliminary X-ray crystallographic analysis of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Mycobacterium tuberculosis*. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **61,** 403–6 (2005).

74. Cornish-Bowden, A. Understanding allosteric and cooperative interactions in enzymes. *FEBS J.* **281,** 621–32 (2014).

75. Lynch, M. Evolutionary diversification of the multimeric states of proteins. *PNAS* **110,** 2821–2828 (2013).

76. Cui, Q. & Karplus, M. Allostery and cooperativity revisited. *Protein Sci.* 1295–1307 (2008).

77. Changeux, J.-P. 50 Years of Allosteric Interactions: the Twists and Turns of the Models. *Nat. Rev. Mol. Cell Biol.* **14,** 819–29 (2013).

78. Monod, J., Wyman, J. & Changeux, J.-P. P. On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **12,** 88–118 (1965).

79. Koshland, D. E., Nemethy, G. & Filmer, D. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* **5,** 365–385 (1966).

80. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allostery. *Nature* **508,** 331–9 (2014).

81. Cooper, A. & Dryden, D. T. F. Allostery without conformational change. *Eur. Biophys. J.* **11,** 103–109 (1984).

82. Tsai, C.-J. & Nussinov, R. A unified view of 'how allostery works'. *PLoS Comput. Biol.* **10,** e1003394 (2014).

83. Kern, D. & Zuiderweg, E. R. The role of dynamics in allosteric regulation. *Curr. Opin. Struct. Biol.* **13,** 748–757 (2003).

84. Del Sol, A., Tsai, C. J., Ma, B. & Nussinov, R. The Origin of Allosteric Functional Modulation: Multiple Pre-existing Pathways. *Structure* **17,** 1042–1050 (2009).

85. Kar, G., Keskin, O., Gursoy, A. & Nussinov, R. Allostery and population shift in drug discovery. *Curr. Opin. Pharmacol.* **10,** 715–722 (2010).

86. Cui, Q. & Karplus, M. Allostery and cooperativity revisited. *Protein Sci.* **17,** 1295–1307 (2008).

87. Smock, R. G. & Gierasch, L. M. Sending signals dynamically. *Science* **324,** 198–203 (2009).

88. Popovych, N., Sun, S., Ebright, R. H. & Kalodimos, C. G. Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **13,** 831–839 (2006).

89. Bahar, I., Chennubhotla, C. & Tobi, D. Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. *Curr. Opin. Struct. Biol.* **17,** 633–640 (2007).

90. Swain, J. F. & Gierasch, L. M. The changing landscape of protein allostery. *Curr. Opin. Struct. Biol.* **16,** 102–108 (2006).

91. Changeux, J.-P. & Edelstein, S. J. Allosteric Mechanisms of Signal Transduction. *Science* **308,** 1424–1428 (2014).

92. Jordan, P. A., Bohle, D. S., Ramilo, C. A. & Evans, J. N. S. New Insights into the Metal Center of 3-Deoxy-D-*arabino*-heptulosonate 7-Phosphate Synthase. *Biochemistry* **40,** 8387–8396 (2001).

93. Blackmore, N. J. The regulation of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Mycobacterium tuberculosis*. *PhD Thesis*

94. Hecker, S. J. & Erion, M. D. Perspective: Prodrugs of Phosphates and Phosphonates. *J. Med. Chem.* 2328–2345 (2008).

95. Nussinov, R. & Tsai, C. Review Allostery in Disease and in Drug Discovery. *Cell* **153,** 293–305 (2013).

96. Fukui, K. Recognition of stereochemical paths by orbital interaction. *Acc. Chem. Res.* **4,** 57–64 (1971).

97. Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J. & Rabinowitz, J. D. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol.* **5,** 593–599 (2009).

98. Subramaniam, P. S., Xie, G., Xia, T. & Jensen, R. A. Substrate Ambiguity of 3-Deoxy-D-*manno*-Octulosonate 8-Phosphate Synthase from *Neisseria gonorrhoeae* in the Context of Its Membership in a Protein Family Containing a Subset of 3-Deoxy-D-*arabino*-Heptulosonate 7-Phosphate Synthases. *J. Bacteriol.* **180,** 119–127 (1998).

99. Li, Y. & Evans, J. N. The hard-soft acid-base principle in enzymatic catalysis: dual reactivity of phosphoenolpyruvate. *PNAS* **93,** 4612–4616 (1996).

100. Heyes, L. C., Reichau, S., Cross, P. J., Jameson, G. B. & Parker, E. J. Structural analysis of substrate-mimicking inhibitors in complex with *Neisseria meningitidis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase - The importance of accommodating the active site water. *Bioorg. Chem.* **57,** 242–50 (2014).

101. Hu, C., Jiang, P., Xu, J., Wu, Y. & Huang, W. Mutation analysis of the feedback inhibition site of phenylalanine-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase of *Escherichia coli*. *J. Basic Microbiol.* **43,** 399–406 (2003).

102. Maeda, H. & Dudareva, N. The shikimate pathway and aromatic amino Acid biosynthesis in plants. *Annu. Rev. Plant Biol.* **63,** 73–105 (2012).

103. Wu, J. & Woodard, R. W. New insights into the evolutionary links relating to the 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase subfamilies. *J. Biol. Chem.* **281,** 4042–4048 (2006).

104. Nazmi, A. R., Schofield, L. R., Dobson, R. C. J., Jameson, G. B. & Parker, E. J. Destabilization of the homotetrameric assembly of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from the hyperthermophile *Pyrococcus furiosus* enhances enzymatic activity. *J. Mol. Biol.* **426,** 656–673 (2014).

105. Swint-Kruse, L. & Brown, C. S. Resmap: Automated representation of macromolecular interfaces as two-dimensional networks. *Bioinformatics* **21,** 3327–3328 (2005).

106. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372,** 774–797 (2007).

107. Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **78,** 1606–1619 (2000).

108. Lebowitz, J., Lewis, M. S. & Schuck, P. Modern analytical ultracentrifugation in protein science : A tutorial review. *Protein Sci.* **11,** 2067–2079 (2002).

109. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25,** 495–503 (1992).

110. Svergun, D. I., Barberato, C. & Koch, M. H. J. CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **28,** 768–773 (1995).

111. Wilkins, M., Gasteiger, E., Bairoch, A., Sanchez, J., Williams, K., Appel, R. & Hochstrasser, D. Protein identification and analysis tools in th ExPASY server. *Methods Mol. Biol.* **112,** 531–552 (1999).

112. Mcphillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Ashley, M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Paul, R., Kuhn, P., Deacon, A. M., Paul, J., Nicholas, K., Phizackerley, R. P., Michael, S., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. & Kuhn, P. Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat.* **9,** 401–406 (2002).

113. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr.* **62,** 72–82 (2006).

114. Evans, P. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr.* **67,** 282–292 (2011).

115. Karplus, P. A. & Diedrichs, K. Linking Crystallographic Model and Data Quality. *Science* **336,** 1030–1033 (2012).

116. McCoy, A. J., Grosse-Kuntslave, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser Crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).

117. Emsley, P., Lohkamp, B., Scott, W. & Cowtan, K. Features and Development of Coot. *Acta Crystallogr.* **66,** 486–501 (2010).

118. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr.* **67,** 355–367 (2011).

119. Murshudov, G. N., Vagin, A. A. & Dodson, E. D. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr.* **53,** 240–255 (1997).

120. Chen, V. B., Arrendall III, B. W., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr.* **66,** 12–21 (2010).

121. Liao, S.-M., Du, Q.-S., Meng, J.-Z., Pang, Z.-W. & Huang, R.-B. The multiple roles of histidine in protein interactions. *Chem. Cent. J.* **7,** 44 (2013).

122. Ma, X., Qi, Y. & Lai, L. Allosteric sites can be identified based on the residue-residue interaction energy difference. *Proteins* 1375–1384 (2014).

123. Stevens, C. Molecular studies of the phanylalanine-inhibited isoenzyme of DAH7PS from *Eschericia coli*. *PhD Thesis* (1991).

124. Mizuguchi, K., Deane, C. M., Blundell, T. L., Johnson, M. S. & Overington, J. P. JOY: protein sequence-structure representation and analysis. *Bioinformatics* **14,** 617–623 (1998).

125. Dam, T. K., Page, D. & Brewer, C. F. Negative Cooperativity Associated with Binding of Multivalent Carbohydrates to Lectins . Thermodynamic Analysis of the ' Multivalency Effect ' †. *Biochemistry* **41,** 1351–1358 (2002).

126. Stevens, S. Y., Sanker, S., Kent, C. & Zuiderweg, E. R. P. Delineation of the allosteric mechanism of a cytidylyl- transferase exhibiting negative cooperativity. *Nat. Struct. Biol.* **8,** 947–952 (2001).

127. Petit, C. M., Zhang, J., Sapienza, P. J., Fuentes, E. J. & Lee, A. L. Hidden dynamic allostery in a PDZ domain. *PNAS* 3–8 (2009).

128. Ferreiro, D. U., Hegler, J. A., Komives, E. A. & Wolynes, P. G. Localizing frustration in native proteins and protein assemblies. *PNAS* **104,** 19819–19824 (2007).

129. Long, D. & Bruschweiler, R. Structural and Entropic Allosteric Signal Transduction Strength via Correlated Motions. *J. Phys. Chem. Lett.* **3,** 1722–1726 (2012).

130. Putcha, B. K., Wright, E., Brunzelle, J. S. & Fernandez, E. J. Structural basis for negative cooperativity within agonist-bound TR : RXR heterodimers. *PNAS* **109,** 6084–6087 (2012).

131. Daily, M. D., Upadhyaya, T. J. & Gray, J. J. Contact rearrangements form coupled networks from local motions in allosteric proteins. *Proteins* **71,** 455–466 (2008).

132. Goodey, N. M. & Benkovic, S. J. Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* **4,** 474–82 (2008).

133. Ricketson, D., Hostick, U., Fang, L., Yamamoto, K. R. & Darimont, B. D. A Conformational Switch in the Ligand-binding Domain Regulates the Dependence of the Glucocorticoid Receptor on Hsp90. *J. Mol. Biol.* **368,** 729–741 (2007).

134. Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent treatment of internal and surface residues in empirical p K a predictions. *J. Chem. Theory Comput.* **7,** 525–537 (2011).

135. Onufriev, A. V & Alexov, E. Protonation and pK changes in protein-ligand binding. *Q. Rev. Biophys.* **46,** 181–209 (2013).

136. Perozzo, R., Folkers, G. & Scapozza, L. Thermodynamics of Protein – Ligand Interactions : History , Presence , and Future Aspects. *J. Recept. Signal Transduct.* **24,** 1–52 (2004).

137. Bembenek, S. D., Tounge, B. A. & Reynolds, C. H. Ligand Efficiency and Fragment-based drug discovery. *Drug Discov. Today* **14,** 278–283 (2009).

138. Le Maréchal, P., Froussios, C., Level, M. & Azerad, R. Enzymic properties of phosphonic analogues od D-erythrose 4 phosphate. *Biochem. Biophys. Res. Commun.* **92,** 1097–1103 (1980).

139. Chen, Y. & Shoichet, B. K. Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.* **5,** 358–64 (2009).

140. Lumley, J. A. Compound Selection and Filtering in Library Design. *QSAR Comb. Sci.* **24,** 1066–1075 (2005).

141. Davis, A. M., Teague, S. J. & Kleywegt, G. J. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed. Engl.* **42,** 2718–36 (2003).

142. Erlanson, D. a, McDowell, R. S. & O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* **47,** 3463–82 (2004).

143. Kobe, A., Caaveiro, J. M. M., Tashiro, S., Kajihara, D., Kikkawa, M., Mitani, T. & Tsumoto, K. Incorporation of rapid thermodynamic data in fragment-based drug discovery. *J. Med. Chem.* **56,** 2155–9 (2013).

144. Larsson, E. A., Jansson, A., Ng, F. M., Then, S. W., Panicker, R., Liu, B., Sangthongpitag, K., Pendharkar, V., Tai, S. J., Dan, C., Ho, S. Y., Cheong, W. W., Poulsen, A., Blanchard, S., Lin, G. R., Alam, J. & Keller, T. H. Fragment-Based Ligand Design of Novel Potent Inhibitors of Tankyrases. *J. Med. Chem.* **56,** 4497–4508 (2013).

145. Meyer, B. & Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angew. Chemie - Int. Ed.* **42,** 864–890 (2003).

146. Dalvit, C., Fogliatto, G., Stewart, A., Veronesi, M. & Stockman, B. WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability. *J. Biomol. NMR* **21,** 349–359 (2001).

147. Mayer, M. & Meyer, B. Characterisation of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chemie - Int. Ed.* **38,** 1784–1788 (1999).

148. Lee, A. L. & Wand, A. J. Microscopic origins of entropy, heat capacity and the glass transition in proteins. *Nature* **411,** 501–504 (2001).

149. Prabhu, N. V & Sharp, K. A. Heat capacity in proteins. *Annu. Rev. Phys. Chem.* **56,** 521–548 (2005).

150. Kovrigin, E. L., Cole, R. & Loria, J. P. Temperature Dependence of the Backbone Dynamics of Ribonuclease A in the Ground State and Bound to the Inhibitor 5 ′ -Phosphothymidine (3 ′ -5 ′) Pyrophosphate. *Biochemistry* **42,** 5279–5291 (2003).

151. Engen, J. R. Analysis of Protein Conformation and Dynamics by Hydrogen / Deuterium Exchange MS. *Anal. Chem.* **81,** 7870–7875 (2009).

152. A. S. Sieben F. J. Simpson, A. S. P. An improved preparative method for D-Erythrose 4-phosphate. *Can. J. Biochem.* **44,** 663–669 (1966).

153. Duke, C. C., Macleod, J. K. & Williams, J. F. Nuclear magnetic resonance studies of D-erythrose 4-phosphate in aqueous solution. Structures of the major contributing monomeric and dimeric forms. *Carbohydr. Res.* **95,** 1–26 (1981).

154. Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. Improved assay for nanomole amounts of inorganic phosphate. *Anal. Biochem.* **100,** 95–97 (1979).

155. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D. Biol. Crystallogr.* **67,** 271–81 (2011).

156. Kabsch, W. Xds. *Acta Crystallogr. D. Biol. Crystallogr.* **66,** 125–32 (2010).

157. Kantardjieff, K. A. & Rupp, B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* **12,** 1865–1871 (2003).

158. Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **33,** 491–497 (1968).

APPENDIX I

General appendices

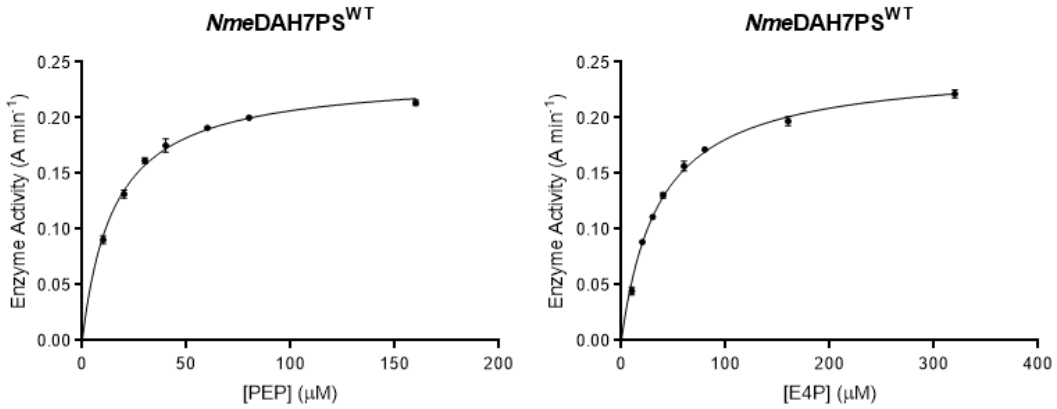


Figure A.1.1 Michaelis-Menten plots of the NmeDAH7PS^{WT} for the K_M of PEP and E4P. Plotted and fitted globally using Graphpad Prism© version 6

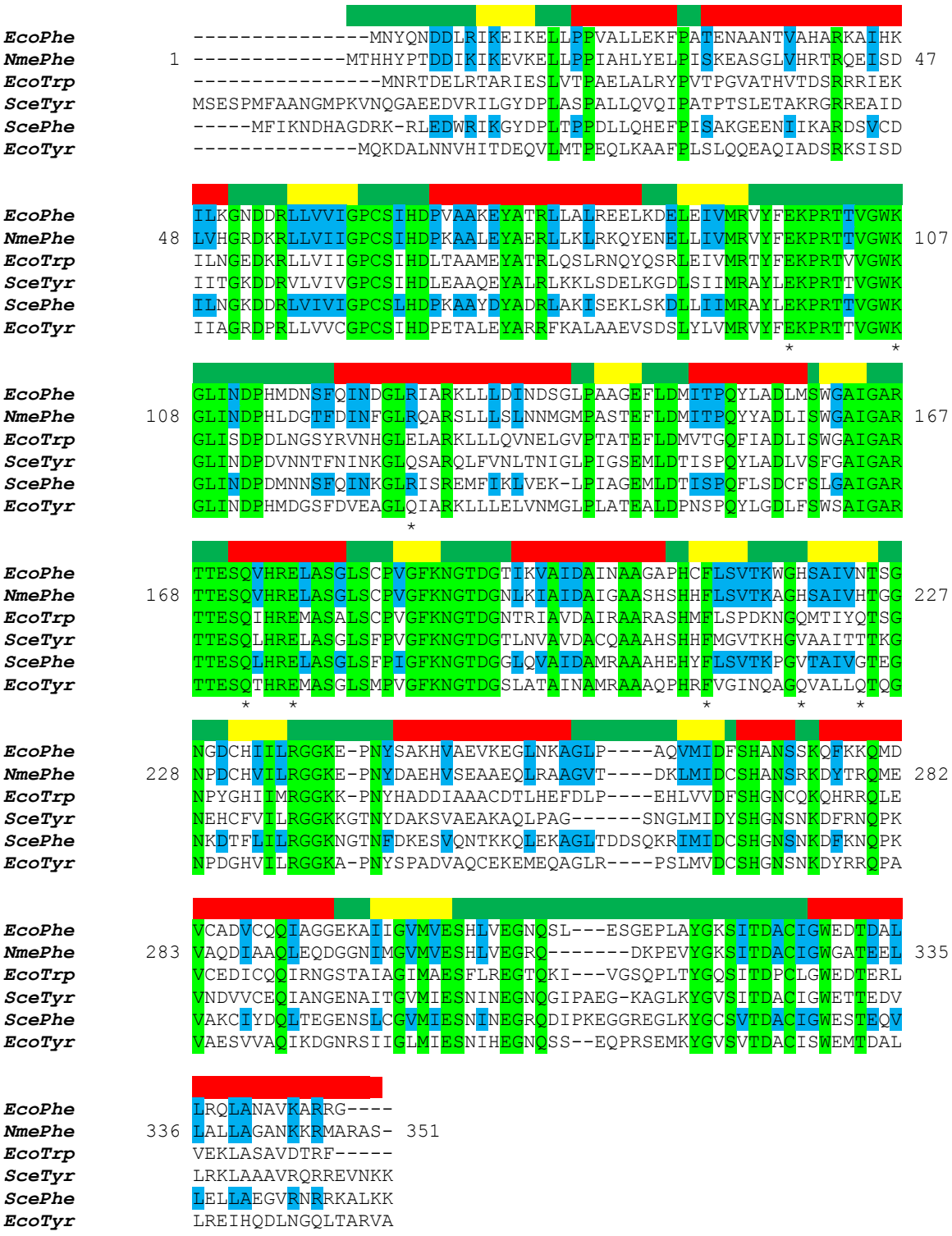


Figure A.1.2 ClustalW sequence alignment of the characterised type Iα DAH7PS isozymes from E. coli and S. cerevisiae in comparison to N. meningitidis DAH7PS. Annotated by NmeDAH7PS helix (red), sheet (yellow) and loop (green). Total conservation across all isozymes (green). Functional conservation across Phe sensitive type Iα DAH7PS isozymes (blue). Mutations involved in this thesis indicated with (*).

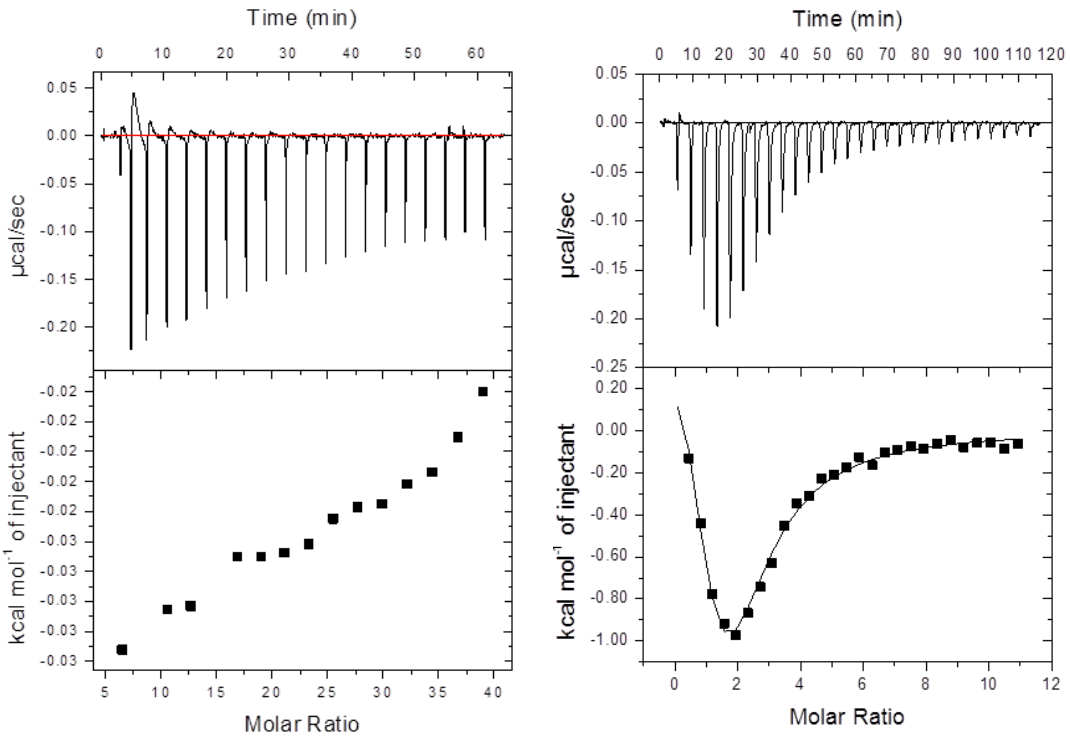


Figure A.3.1 Left. Raw data of the ITC experiments for the binding of Trp (20mM) to NmeDAH7PS^{WT} (100µM) using the ITC200 Microcal™.Right. Raw data of the ITC experiments for the binding of Phe (0.8mM) to NmeDAH7PS^{WT} (16µM) using the VP-ITC Microcal™

Chapter 4

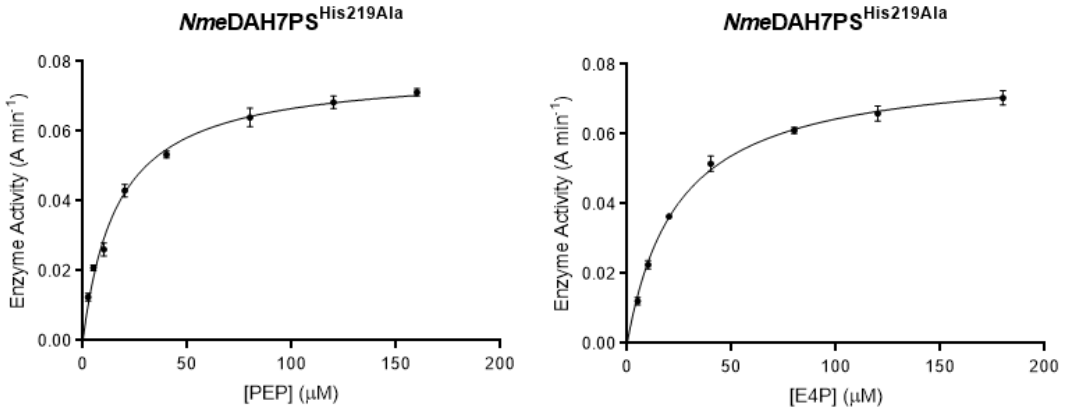


Figure A.4.1 Michaelis-Menten plots of the NmeDAH7PS^{His219Ala} for the K_M of PEP and E4P. Plotted and fitted globally using Graphpad Prism© version 6

Chapter 5

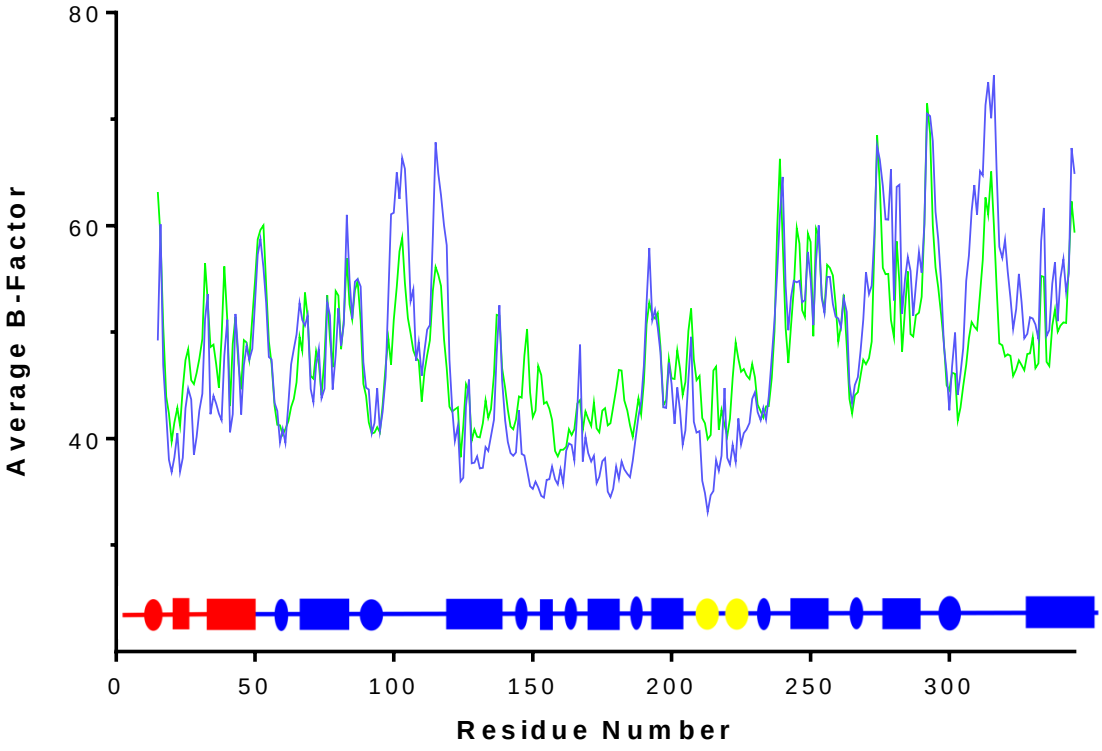


Figure A.5.1 Normalised average B-factors for all chains of the Phe-bound (blue) and PEP-bound (green) DAH7PS (PDB codes: 4UC5 and 4HSN).

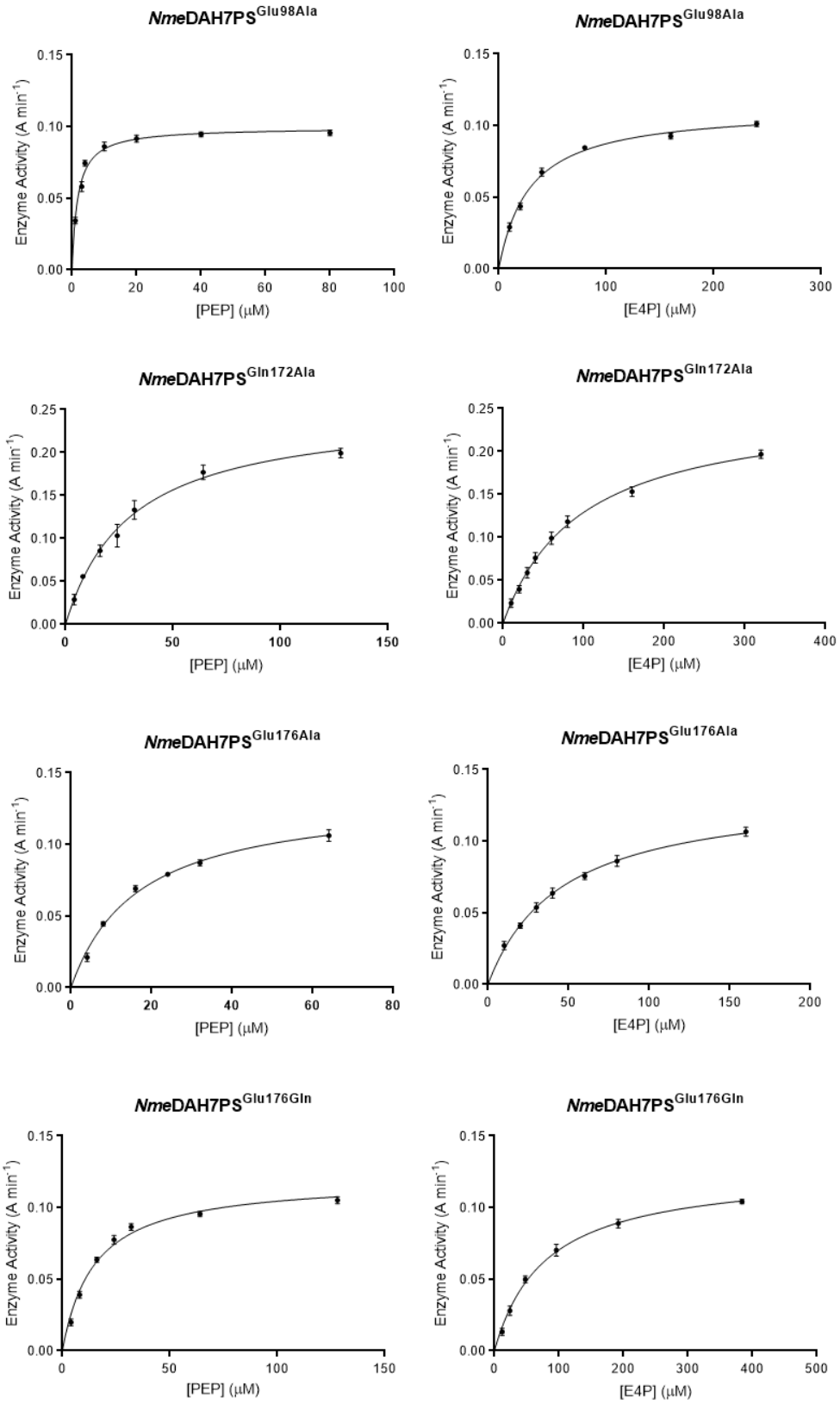


Figure A.5.2 Michaelis-Menten plots of the NmeDAH7PS^{Glu98Ala}, NmeDAH7PS^{Gln172Ala}, NmeDAH7PS^{Glu176Ala} and NmeDAH7PS^{Glu176Gln} for the K_M of PEP and E4P. Plotted and fitted globally using Graphpad Prism© version 6

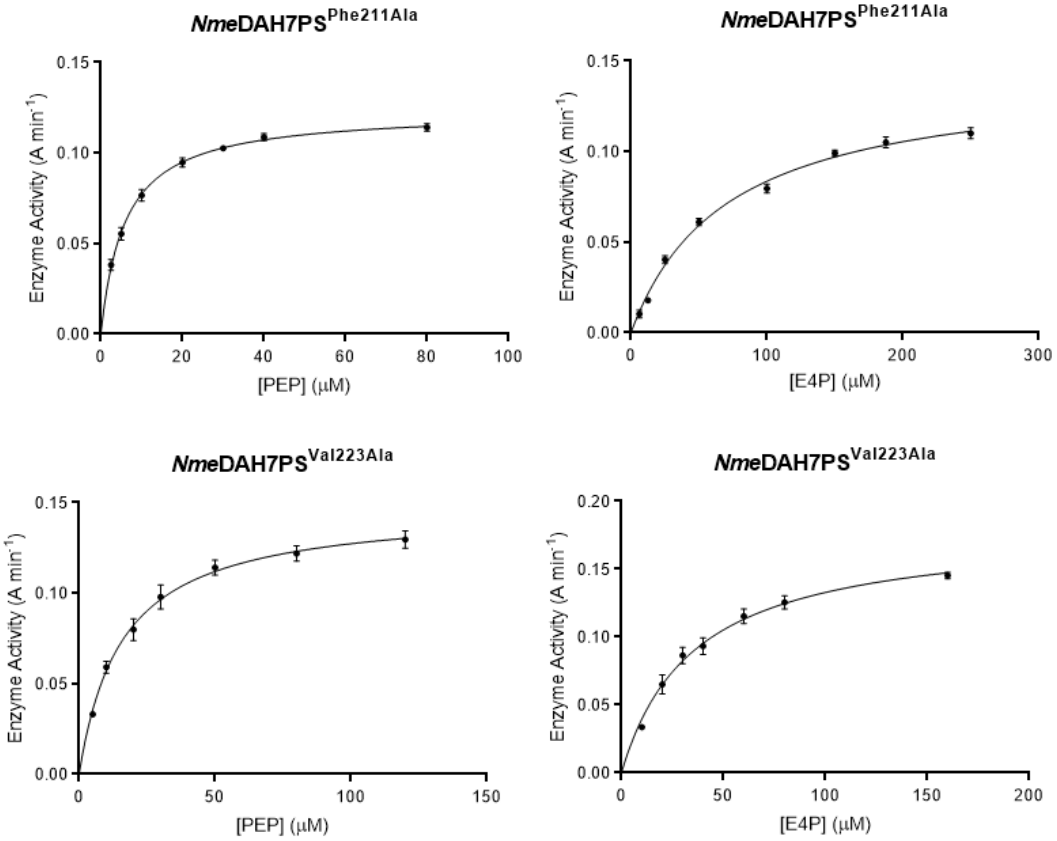


Figure A.5.3 Michaelis-Menten plots of the NmeDAH7PS^{Phe211Ala} and NmeDAH7PS^{Val223Ala} for the K_M of PEP and E4P. Plotted and fitted globally using Graphpad Prism© version 6

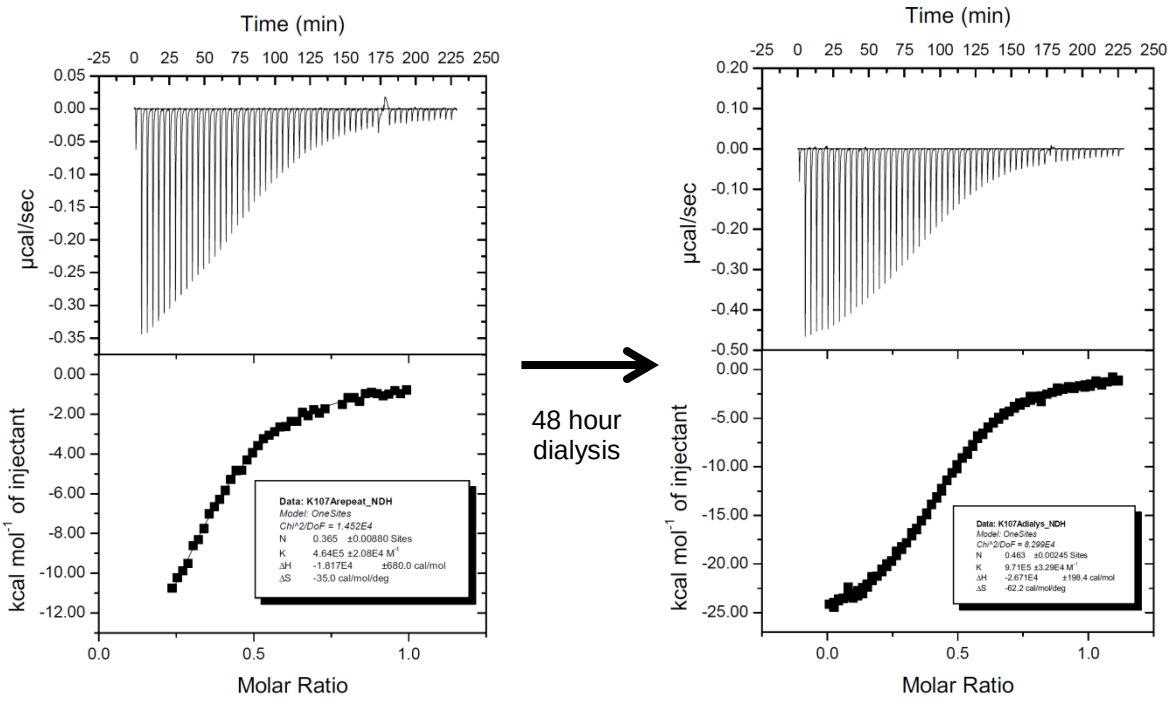


Figure A.5.4 Binding isotherms of Phe to NmeDAH7PS^{Lys107Ala} before and after 48 hours of dialysis.

Calculated pK_A variations expose dynamic allosteric communication networks.

Eric J M Lang[a‡], Logan C Heyes[a‡], Geoffrey B Jameson[b] and Emily J Parker[c*]

[a] Biomolecular Interaction Centre and Department of Chemistry, University of Canterbury, PO Box 4800, Christchurch 8140, New Zealand. [b] Institute of Fundamental Sciences, Massey University, PO Box 11-222, Palmerston North 4422, New Zealand. [c] Maurice Wilkins Centre, Department of Chemistry, University of Canterbury, PO Box 4800, Christchurch 8140, New Zealand

ABSTRACT: Allosteric regulation of protein function, the process by which binding of an effector molecule provokes a functional response from a distal site, is critical for metabolic pathways. Yet, the way the allosteric signal is communicated remains elusive, especially in dynamic, entropically-driven regulation mechanisms for which no major conformational changes are observed. To identify these dynamic allosteric communication networks, we have developed an approach that monitors the pK_a variations of ionizable residues over the course of molecular dynamics simulations performed in the presence and absence of an allosteric regulator. As the pK_a of ionizable residues depends on their environment, it represents a simple metric to monitor changes in several complex factors induced by binding an allosteric effector. These factors include coulombic interactions, hydrogen bounding and solvation, as well as backbone motions and sidechain fluctuations. The predictions that can be made with this method concerning the roles of ionizable residues for allosteric communication can then be easily tested experimentally by changing the working pH of the protein or performing single point mutations. To demonstrate the method's validity, we have applied this approach to the subtle dynamic regulation mechanism observed for *Neisseria meningitidis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase, the first enzyme of aromatic biosynthesis. We were able to identify key communication pathways linking the allosteric binding site to the active site of the enzyme and to validate these findings experimentally by reestablishing the catalytic activity of allosterically inhibited enzyme via modulation of the working pH, without compromising the binding affinity of the allosteric regulator.

INTRODUCTION

Allosteric regulation of protein function is critical to cellular metabolic control. Allostery is the process by which the binding of a ligand provides a functional response at a distant site, and is essential for cell signaling events and sophisticated metabolic control through the regulation of enzymatic catalysis. Traditional models of allostery described the allosteric response as being mediated by large conformational changes that directly affect activity.[1,2] However, using the formalism of statistical mechanics, Cooper and Dryden [3] proposed that allostery could exist even in the absence of conformational changes via a purely dynamic, entropically driven, mechanism. Recent advances in protein NMR, molecular dynamics (MD) simulations and isothermal titration calorimetry (ITC) have illustrated that subtle changes in dynamic fluctuations of proteins can govern the communication of allosteric information between distant sites without conformational change, and that even intrinsic disorder and local unfolding can play an important role in allosteric regulation, thereby confirming the role of dynamics in allostery ([4] and references therein).

This concept of dynamic allostery is receiving increasing recent attention.[4-6] However, understanding of how the allosteric signal is communicated at an atomic level between distant sites remains challenging, due to the subtle, transient and complex changes in dynamics encountered during entropically driven allosteric regulation. It is now well accepted that the allosteric signal is propagated via multiple pathways through a network of physically and/or thermodynamically interconnected residues.[7-9] Allosteric ligand binding perturbs this communication network and results in a functional response by the protein.

However, elucidating such communication networks remains elusive, as they cannot be readily recognized by large changes in residue interactions and may include entropic (hydrophobic interactions and changes in protein flexibility) as well as enthalpic (van der Waals, hydrogen bonding and coulombic interactions) effects. Hence, the most recent approaches take advantage of MD simulations and aim to identify and characterize these allosteric networks[10] using correlation analysis of residue-residue contacts,[11] dynamical network models,[12,13] force distribution analysis,[14,15] degree of frustration,[16] evolutionary covariance,[17] or correlated motions using mutual information theory.[18,19] These approaches provide compli-

cated communication networks that often need to be coarse-grained for analysis, leading to the loss of atomic-level details. This results in the omission of electrostatic effects during analysis despite their potential influence.[10] Moreover, experimental validation of these reconstructed networks is uncommon, due to their complexity.

Here we propose and test an approach that relies on the analysis of pK_a variations of ionizable residues during MD simulations as a means of capturing the allosteric communication pathway. The pK_a of ionizable residues are neither fixed nor constant in time, but rather depend on their local environment and dynamics.[20] Therefore substantial change in the time-averaged pK_a of a given residue upon binding an allosteric ligand suggests that this particular residue or its neighbors may play a role in the allosteric regulation of the enzyme. This approach has the marked advantage of taking into account backbone motions, side-chain fluctuations, solvent exposure, hydrogen bonding and other electrostatic interactions, which are otherwise difficult to consider individually. The resulting global pK_a metric can be easily deconvoluted to account for the contribution of individual components, and to provide a means to probe dynamic allostery. This approach leads to experimentally verifiable predictions due to the possibility of perturbing a communication pathway by either mutating an ionizable residue or changing the working pH.

To demonstrate the benefits of such an approach, we have used this method to shed light on the allosteric communication pathway of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from the pathogen *Neisseria meningitidis* (*Nme*DAH7PS). This enzyme catalyzes the first committed step of the shikimate pathway in bacteria, responsible for aromatic amino acid biosynthesis, and is allosterically regulated by the pathway end-product phenylalanine.[21] Its mechanism of allosteric response is particularly challenging to delineate as the enzyme does not experience any major conformational change in its regulated state. Instead, regulation appears to be mediated by subtle fluctuations in dynamics upon binding of the allosteric effector, phenylalanine (Phe). By tracking the pK_a changes and fluctuations associated with this allosteric effector, our study illuminates a dynamic network of non-covalent interactions that affects primarily the dynamics of the protein to ultimately regulate its catalytic activity.

RESULTS

Comparison of Phe-bound and Phe-free structures implicates a hydrogen-bonding network in the allosteric response. DAH7PS is a metalloenzyme that catalyzes the condensation of erythrose 4-phosphate (E4P) and phosphoenolpyruvate (PEP) to form DAH7P.[21] PEP binds deep within the active site, located at the C-terminal ends of the beta strands of the $(\beta/\alpha)_8$ barrel, and E4P interacts principally with the residues of the $\beta_2\alpha_2$ loop (Figure 1). *Nme*DAH7PS is homotetrameric, consisting of two pairs of active, regulated dimers[22] that share an extensive interface. There is a single allosteric site per chain, located close to the dimer interface, and the pair of sites is separated by a distance of ~25 Å. The allosteric site is located ~24 Å from its nearest active site, which is on the same chain, and is ~30 Å across the dimer interface from the active site of the opposing chain. The allosteric site is formed in part by an N-terminal extension from the opposing chain, which closes over the allosteric ligand, Phe, and also by the β-hairpin (β_{6a} and β_{6b}) insertion into the $\alpha_5\beta_6$ loop that contributes to the dimer interface (Figure 1A).

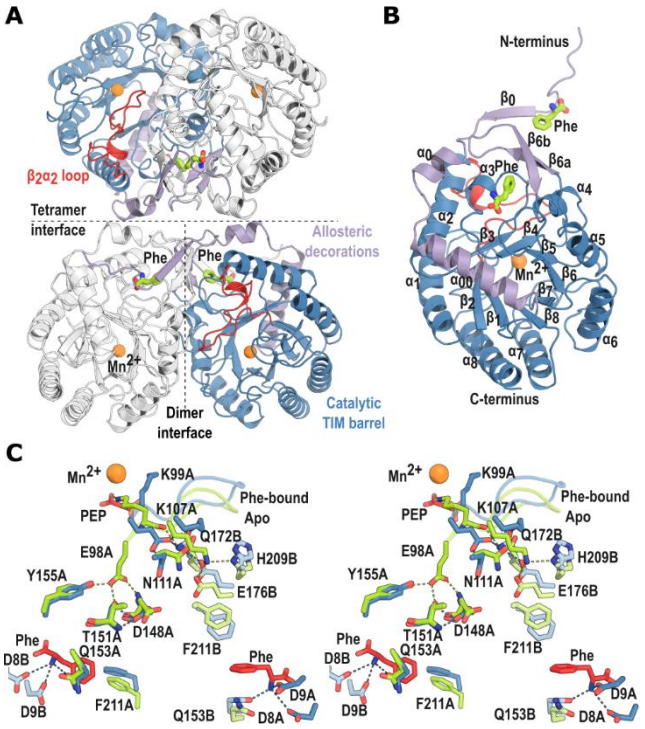


Figure 1. A. Homotetrameric Phe-bound *Nme*DAH7PS, for each dimer, one chain is displayed in white while the adjacent chain is colored with the following scheme: main $(\beta/\alpha)_8$ barrel in blue, allosteric decorations in purple and catalytically important $\beta_2\alpha_2$ loop in red. Mn^{2+} is shown as an orange spheres and Phe as green sticks in each monomeric unit. **B.** Monomeric unit of Phe-bound *Nme*DAH7PS with using the same color scheme as describe above and with the name of the secondary structure elements displayed **C.** Stereoview of the overlay of the dimer interface of chain A and B of *Nme*DAH7PS for the apo (green) and Phe-bound (blue) forms. Hydrogen bonds are colored with respect to the structure they originate; apo, green and regulated, blue. Maps showing electron density for bound Phe are shown in Figure S1.

A crystal structure of *Nme*DAH7PS bound to allosteric inhibitor Phe was determined at 2.2 Å (PDB code 4UC5). Comparison of this structure with the unliganded struc-

ture (PDB code 4HSN)[21] gives a root-mean-square deviation (RMSD) of 0.96 Å for the alignment of 1302 Cα atoms of the tetramer, and 0.583-0.752 Å for the alignment of 639-641 Cα atoms from all dimer pairings. There is a ~2° twist of one dimer relative to the other in the presence of Phe accounting for the larger change observed when the tetramers are compared. Despite the overall similarity between Phe-free and Phe-bound structures, it is evident that areas of the protein are significantly displaced in the Phe-bound structure. In particular, the N-terminal extension (residues 1-10) becomes ordered around the bound Phe. Additionally, the $\beta_2\alpha_2$ loop (residues 95-120) is repositioned, with an average Cα RMSD from the apo structure of 2.3 Å. This change is associated with some differences in hydrogen-bonding partners for residues of this loop. Hydrogen bonds between Lys107 and residues Glu176, His209, and the backbone carbonyl of His210 of the opposing chain are lost, whereas new hydrogen bonds are established between Lys107 and Glu98 and Gln172 of the opposing chain. Likewise, hydrogen bonds between Glu98 and residues Thr151, Tyr155 and the backbone of Asp148 are lost, as is the hydrogen bond that links the backbone carbonyl of the catalytic Lys99[23] with Gln172 of the opposing chain.

Allosteric regulation triggers local flexibility changes and leads to a loss of correlated motions. To study the dynamics of *Nme*DAH7PS in solution, molecular dynamics (MD) simulations of both the apo and Phe-bound forms of the tetrameric enzyme were conducted in explicit water after modelling of the missing residues from both structures. For each system two simulation runs of 200 ns were performed starting from slightly different structures (see Methods). As it has recently been demonstrated that the dimer is the elementary biological unit of the enzyme, capable of both catalysis and regulation,[22] analysis of the trajectories has been carried out on the dimeric unit, leading to a combined total of 800 ns simulated for both the apo and Phe-bound dimers.

The root-mean-square deviations (RMSD) of the alpha carbons along the trajectories indicate that the simulations can be considered as fully equilibrated after about 80 ns, therefore, the following analyses are conducted on the last 120 ns of each trajectory. The dimers in the presence and absence of Phe sample slightly different, non-overlapping conformations but do not deviate from their respective crystal structure by more than 3.7 Å (Figure 2A). Overall, the RMSD results indicate that although the active and inhibited forms present a slightly different dynamic profile, both do not undergo large conformational changes or domain motions within the period of analysis and the differences between the apo and Phe-bound states are subtle.

The distribution of the distances between the center of mass of each monomeric unit composing the functional dimer is more narrow in the absence of Phe and has a lower average calculated distance, suggesting a slight increase in dynamics for each chain with respect to the other and a looser association of subunits to form the dimer in the presence of Phe (Figure S2).

Alpha carbon root-mean-square fluctuations (RMSF) averaged per chain and per run, reveal only limited differences in the flexibility of the apo and Phe-bound chains, confirming the absence of major conformational changes over the course of the simulation (Figure 2B). Unsurprisingly, the flexibility of the N-terminal tail, which covers the allosteric site upon binding of Phe, is greater for the apo form. This is also the case for the $\alpha_o\alpha_{oo}$ loop which comes in contact with the tip of the N-terminal tail in the presence of Phe. Apart from the N-terminal region, the greatest change in observed upon binding of Phe corresponds to key catalytic $\beta_2\alpha_2$ loop, which becomes significantly more flexible in the presence of the allosteric ligand. In addition, binding of Phe leads to an increase in flexibility of the $\beta_3\alpha_3$ loop, β-hairpin insertion and, the region between residues 150 and 180 (β_4 to α_4 region). Finally, the external helix α_7 appears to experience greater fluctuation in the Phe-bound form while the $\beta_8\alpha_8$ loop becomes more rigid.

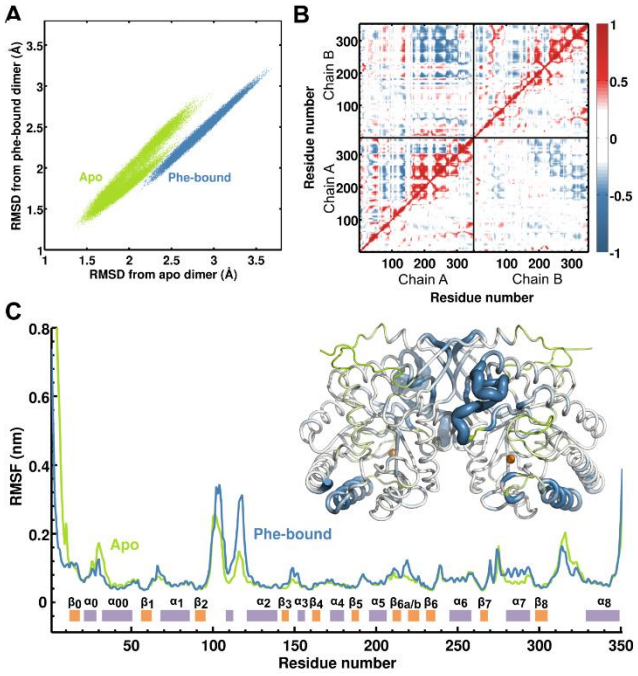


Figure 2. A. RMSD of the alpha carbons of the dimer along the last 120 ns of simulation of each MD runs in the presence (blue dots), absence (green dots) of allosteric Phe, calculated using the X-ray crystal structure of the apo (x coordinates) and Phe-bound (y coordinates) forms as references. **B.** Correlations between alpha carbons of the dimer calculated using the Pearson coefficient. The upper left triangle corresponds to the apo form and the lower right triangle to the Phe-bound form. Correlation values are represented by a color gradient of blue to red. Correlated motions with an absolute value below 0.25 are neglected. **C.** Cα root-mean-square fluctuations (RMSF) of the apo (green) and Phe-bound (blue). Secondary structure of the monomeric unit is depicted with

orange and purple rectangles to identify β-strands and α-helices respectively. The difference in RMSF between the Phe-bound and apo forms is displayed on the dimer structure. Regions which are more flexible in the presence of Phe are represented with an increased thickness and a blue color gradient while regions more flexible in the apo trajectory are represented with decreased tube diameter and a green color gradient.

Overall, both intra- and inter-chain motions between residues are more correlated in the absence of allosteric inhibitor, as revealed by the normalized matrix of correlated motions between residues obtained from covariance matrix of the alpha carbon atoms of each residue. This suggests that the changes associated with the binding of Phe perturb the network of non-bonding interactions, resulting in a lower degree of communication between residues and more random motions, consistent with an entropically-driven allosteric mechanism.

pK_a analysis identifies ionizable residues involved in the allosteric response. The calculated pK_a values for the ionizable residues of each frame of the apo and Phe-bound trajectories were determined using PROPKA3.1.[24] By taking into consideration all the conformations sampled during the MD simulations, this approach provides a robust means to assess the average pK_a value of a particular residue, how much it fluctuates around this average, and the factors (desolvation effects, hydrogen bonds and coulombic interactions) influencing this value.

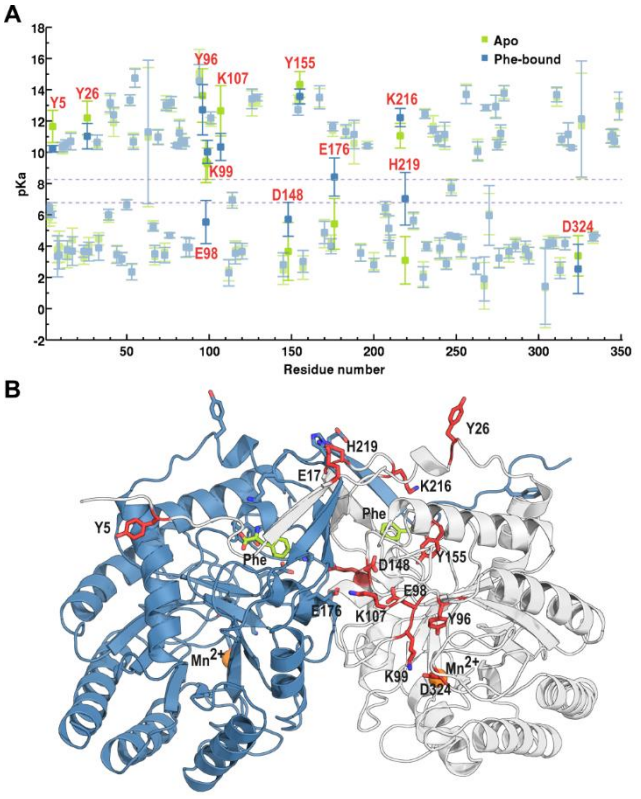


Figure 3. A. The average pK_a values for the apo and Phe-bound forms are represented with green and blue squares respectively (also presented in Figure S3), while the error bars correspond to ± one standard deviation around the average (also presented in Figure S4). The lowest (6.8) and highest (8.3) experimental pH used are represented with dotted purple lines. Ionizable residues that present a major pK_a shift upon binding of Phe have their residue number indicated in red while the other ionizable residues appear faded in the background. **B.** Structure of the dimer (one chain is shown in blue, the other in white) with residues experiencing major pK_a shift represented in sticks. The allosteric site is highlighted by the presence of Phe shown in green sticks, while the catalytic site is identified with the manganese ions shown as orange spheres.

The differences between the pK_a of the ionizable residues of the apo and Phe-bound forms, averaged over the course of the MD simulations, enable a number of residues likely to be involved in the allosteric communication pathway to be identified. Twelve residues were calculated to experience a significant pK_a shift upon binding of Phe (Figure 3A). These residues are found to be at key positions or to have a specific role (Figure 3B): Lys216 is involved in the binding of allosteric Phe and Lys99 is responsible for proton-transfer reaction within the active site as part of catalysis. Glu98 and Lys107 are found on the $\beta_2\alpha_2$ loop and interact with Asp148, whereas Glu176 interacts with Lys107 across the dimer interface in the apo form. As the backbones of residues Asp148, Tyr155 and Glu176 are relatively close to the allosteric binding sites, they are likely to be affected by the binding of Phe. Tyr26 and His219 are positioned at the tetrameric interface. The predicted pK_a shift of Tyr5 is likely to be due to the alteration in the flexibility of the N-terminal tail.

The standard deviation of the pK_a values of each ionizable residue over the course of the MD trajectories gives an indication of how much its pK_a value is predicted to fluctuate in the presence or absence of Phe (Figure 3A). These fluctuations may be associated with changes in conformational dynamics for the residue itself or for its surrounding environment. It is also apparent that a shift in the predicted average pK_a value upon binding of Phe is not necessarily associated with a change in calculated pK_a fluctuation, with Glu98 being the best example of such behavior. This observation implies that although the environment of the residue changes in the presence of Phe, this residue remains tightly bound and/or buried in both inhibited and active forms. The opposite behavior is more common; several residues experience a change in the fluctuation of their pK_a value upon binding of Phe, but do not experience any major pK_a shift of their average value. On the other hand, Lys107 experiences a drop in pK_a fluctuations, as well as in pK_a itself, on the binding of Phe.

Interestingly, of the twelve residues for which a major pK_a shift is predicted, Tyr5, Glu98, Lys107, Glu176 and His219 present (almost) no overlap between their values at ± one standard deviation (Figure 3A), supporting the

existence of a significant pK_a difference between the apo and Phe-bound forms. On the other hand, changes in average desolvation are also predicted for Glu98, Lys99 and Lys107 (Figure S5), which experience an increased solvent exposure in the presence of allosteric Phe and for Tyr155, as well as Asp8 and Arg42, which bind allosteric Phe and which become more buried in the presence of Phe.

Inter- and intra-chain communication changes upon allosteric inhibitor binding. To identify changes in the hydrogen bonds and coulombic interactions between pairs of residues upon binding of Phe that are likely to affect the pK_a of ionizable residues, a weighted map of interactions in which two nodes represent two interacting residues and the diameter of the edge linking two nodes accounts for the strength of the interaction was generated. By specifying the working pH, we can predict which residues are protonated and thus account for the favorable, unfavorable or non-existent coulombic interactions depending on the charge borne by the residues (Figure 4A).

The hydrogen bonding and coulombic interaction maps reveal clearly that in the absence of allosteric inhibitor, the

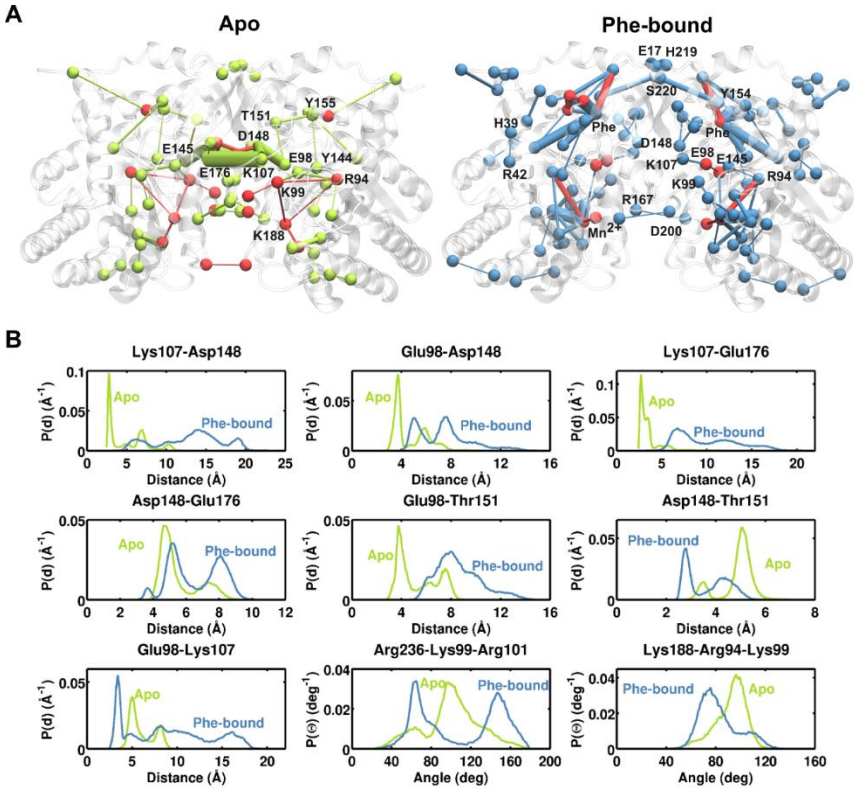


Figure 4. A. Hydrogen bonds and coulombic interactions that are only present or are stronger in the apo (top) or the Phe-bound (bottom) trajectories. Nodes represent interacting residues and the edges between nodes indicate the interactions, with the diameter of the edge accounting for the strength of the interaction. For the apo form, favorable interactions are shown in green and unfavorable ones are in red, and in the Phe-bound form favorable interactions are in blue and unfavorable ones are in red. The contribution of each effect (side chain hydrogen-bond, backbone hydrogen-bond and coulombic interactions) is detailed in FigureS6. **B.** Distribution of the distances between key ionizable groups. Frequency distribution of the distances for the apo (green) and Phe-bound (blue) forms. The distances are calculated using the hydroxyl oxygen of Ser, Tyr, Thr the center of the two carboxyl oxygens of Glu and Asp, the nitrogen of the ε-amino group of Lys, the center of the guanidinium ion of Arg and the center of the two nitrogens of the imidazole ring of His.

key predicted interactions are the favorable coulombic interactions between Glu176 and Lys107 across the dimer interface, and between Lys107 and Asp148; the unfavorable ones are between Asp148 and Glu176 across the dimer interface, as well as the hydrogen bond between Asp148 and protonated Glu98 (Figure 4A). These strong interactions ensure close contact between the two chains of the core dimeric unit of the enzyme and provide a short and robust inter-chain communication pathway between the allosteric site of one chain and the active site of the adjacent chain in the absence of Phe. Upon binding of allosteric Phe all these strong interactions are lost, essentially leading to the disruption of the network of interactions between ionizable residues that exists across the two

chains (Figure 4A). This observation is confirmed by examining the calculated distance distributions between the ionizable groups of these residues over the course of the MD simulations (Figure 4B). The calculated distance of 2.7 Å between Lys107 and Asp148 in the apo form is shifted to 6 - 19 Å in the presence of Phe, whereas the distance between Glu98 and Asp148 changes from 3.7 Å in the apo form to around 5 Å. The distance between Glu176 and Lys107 across the tight dimer interface is mainly centered around 2.7 and 3.5 Å in the absence of ligand, but is not closer than 5 Å and is calculated to reach as much as 18 Å in the presence of the allosteric ligand. Finally the distance between the two negatively charged residues Glu176 and Asp148 in the apo form across the interface remains centered around 4.6 Å (Figure 4B), whereas binding of Phe and protonation of Glu176 enables this distance to reduce to as little as 3.6 Å between the two ionizable groups, although the distances are mainly centered around 5.2 and 8 Å.

Within a single chain of the unregulated enzyme, Tyr155, which is close to the allosteric site, interacts via a hydrogen bond with Thr144 (adjacent to catalytically important Glu145) and with Thr151, which in turn interacts with Glu98. These interactions, which are lost upon binding of Phe, may contribute to communicate the allosteric information within a single chain. In the presence of Phe, the distance between Glu98 and Thr151 increases, while the distance between Thr151 and Asp148 decreases with the pair interacting via a hydrogen bond (Figure 4B). Binding of Phe also leads to the formation of strong interactions within and around the allosteric site, which tighten the upper part of the dimer.

A number of important changes also occur within and around the active site. In the apo form, repulsive coulombic interactions, exist between catalytically important residues Lys188, Lys99, both of which are involved in the proton-transfer mechanisms during catalysis, Arg94 which binds the carboxylate of PEP, as well as Arg236 and Arg101 which bind the phosphate of PEP and E4P respectively[23]. This network of repulsive interactions, which is only present in the absence of Phe, between positively charged residues is likely to maintain the correct relative distances between key residues, thereby maintaining the most catalytically active conformation. Additionally, Glu145 interacts more favorably with Lys99 and Arg94 in the absence of Phe. Binding of Phe reshuffles this interaction network, whereby, Lys107 and Glu98 now interact via a hydrogen bond and move closer to one another with a distance of 3.4 Å, compared with 5 Å in the absence of ligand (Figure 4B). The negative charge on Glu98 has an unfavorable effect on Glu145 which interacts less with Arg94 and Lys99 with the functional groups of these two residues moving away from Glu145. The slight balance of repulsive charge-charge interactions between catalytically important Lys99, Lys188 and Arg94 is also broken and the distances between the functional groups of these residues increase. Significant changes in the angle formed by the functional groups of these interacting residues also support a shift from an active to an inactive topology of the key catalytic residues upon binding of Phe (Figure S7). Finally a large number of existing interactions centered on the metal ion are significantly strengthened in the presence of Phe.

The inhibitory effect of allosteric Phe is reduced at elevated pH. The pK_a calculations of the regulated and apo forms of *Nme*DAH7PS exposed several residues that likely adopt significantly altered pK_a values in the presence and absence of Phe (Figure 5). Of particular interest are Glu98 and Glu176, which both present unusually high pK_a values for Glu and are predicted to have different protonation states at standard assay pH depending on the presence of Phe. At an assay pH of 6.8 Glu98 is predicted to be protonated in the apo enzyme and deprotonated in the Phe-bound form, whereas Glu176 is predicted to be deprotonated in the apo form, but protonated in the presence of Phe. As the pK_a values of these residues may lie close to assay pH, changing their protonation state experimentally through variations in pH may mimic in part the allosteric response. In the absence of Phe, over the pH range 6.8 to 8.3 there is no significant change in the protonation state of these two residues and little effect of pH on activity is predicted or observed (Figure 5A).

To investigate the pH dependence of the allosteric response, assays were conducted between pH 6.8 and 8.3 in the presence of 300 μM Phe, as enzyme activity decreases only slightly over this pH range (Figure 5A). In addition, outside this range, the protonation states of the substrates and important catalytic residues may be affected. The only change in protonation state predicted from the pK_a analysis with increasing pH over this range is that Glu176 becomes substantially protonated by pH 8.3 in the presence of Phe.

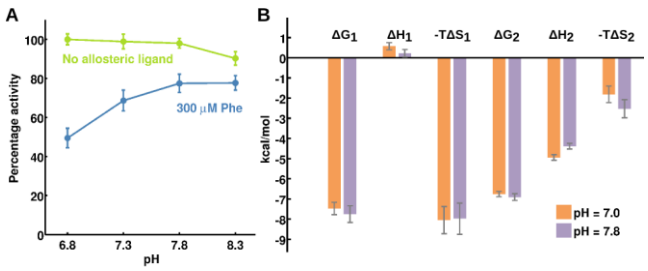


Figure 5. A. The effect of pH on the activity and inhibition of *Nme*DAH7PS in the presence of 300 μM L-phenylalanine. Activity at varying pH compared to maximal activity seen at pH 6.8 (green line). Activity at varying pH in the presence of 300 μM Phe compared to uninhibited enzyme at pH 6.8 (blue line). Substrate concentrations were at least 5-fold higher than the K_m values (assessed at the different pH values) of each substrate at each pH (substrate concentrations were 100 μM PEP and 500 μM E4P). **B.** The entropic and enthalpic contributions to the binding isotherm of Phe to *Nme*DAH7PS at pH 7 (orange) and pH 7.8 (purple).

The specific activity was observed to decrease slightly between pH 6.8 and 7.8, with a sharper reduction at pH 8.3 (to 90% maximal activity, Figure 5A). More importantly, in the presence of Phe the inhibitory response was observed to be severely attenuated as the pH was raised (Fig 5A.), consistent with the dependence of *Nme*DAH7PS allostery on the network of ionizable residues for signal transmission. Comparable K_D values were measured for the binding of Phe at both pH 7 and pH 7.8 (the maximum pH for enzyme stability over the course of the experiment), indicating that the binding of Phe is not compromised in this pH range. Moreover, the entropic and enthalpic contributions to the free energies of binding of Phe to each site in the sequential two-site binding model are very similar at both pH values (Figure 5B).

DISCUSSION

Evolution has favored the existence of subtle yet fundamental, dynamic networks of non-covalent interactions within proteins that are intimately related to function. These non-covalent interactions consist primarily of hydrogen bounds and coulombic interactions but also involve hydrophobic interactions, water-mediated hydrogen bonding. As a protein samples its energy landscape, interactions between residues that comprise the network are constantly changing. A method of altering the dynamics of a protein and therefore its function is to alter the existing network by favoring transient, non-covalent interactions through the binding of an allosteric effector molecule in order to reshape the conformational space sampled by the protein. This altered state will manifest as the product of the interactions broken and formed in the network linking the two remote sites. In order for the information from a remote site to be conferred to the active site, communication through these complex non-covalent bonding networks is essential.

Allosteric regulation of *Nme*DAH7PS relies on changes in the intricate network of interactions that connect allosteric and catalytic sites of the protein, both within a single chain and across the dimer interface. This network undergoes a significant reorganization to evoke the allosteric regulatory response in the enzyme. For the apo enzyme this network enables the communication between the allosteric site and the catalytic site through the dimer interface in particular via strong coulombic forces among Glu98, Lys107, Asp148 and Glu176, but also within a single chain via hydrogen bonds involving Thr151, Tyr155, Glu145 and Glu98. The presence of the allosteric effector molecule Phe, leads to the loss of these two communication pathways between allosteric and active sites and an increase of non-correlated motions within the enzyme. The catalytic activity of the enzyme appears to be considerably lowered mainly because of the increased flexibility of the $\beta_2\alpha_2$ loop, providing an entropic penalty to substrate binding, and disruption of the optimal, finely tuned, relative position of key catalytic residues, providing a primarily enthalpic penalty to substrate binding.

Identification of the changes occurring in the reorganization of this network is revealed by monitoring the calculated pK_a variability between the active and regulated forms during the course of an MD simulation. Previous similar studies have, however, been made on a limited selection of structures generated from MD to depict protein flexibility[25-32]. Our studies show that the use of comprehensive sampling of conformational space leads to the complete recovery of the vast information pK_a calculations can provide and enable the formulation of robust conclusions based on the observed variations.

In principle, constant-pH molecular dynamics (CpHMD)[33-35] should be the method of choice for studying pK_a variations as a function of protein motions. However, such an approach cannot be applied as an analytical tool on existing trajectories and is usually associated with a computationally expensive replica-exchange method that cannot be applied to large multimeric proteins. Apart from CpHMD, the pK_a of ionisable residues in a protein can be assessed using either methods that relies on the Poisson-Boltzmann[36] or Generalized Born[37] continuum models (which are for example implemented in MEAD,[38] MCCE,[39] pkD,[40] UHBD[41] and H++[42]); or empirical methods such as implemented in PROPKA.[24] PROPKA has been shown to outperform electrostatic based methods[27] while reducing computational cost and is thus well suited for performing pK_a analysis of complete, existing MD trajectories.

It should be emphasized that the key residues that are predicted to play an important role in the signal transmission are not readily identified by comparing the pK_a or structural changes of the crystal structures. Indeed, not surprisingly such an approach is precluded by the inherent static image of the interactions, which leads to a number of 'false positives' and to the oversight of key interactions. Combining structural analysis, MD simulations and pK_a analysis of the MD trajectories therefore greatly improves the robustness of the results and enables to refine the interpretation of the findings to lead experimentally verifiable results.

Changes in the pK_a of ionizable residues result from distinct changes in the non-covalent interaction network connecting the functional sites of *Nme*DAH7PS. We have shown that alteration of the protonation state of these crucial ionizable residues by altering the pH, limits the ability of the enzyme to effectively reorganize its interaction network in the presence of Phe.

The combination of MD and pK_a calculations exposes the communication network between functionally important remote sites. The methodology described here to illustrate the complexities of 'dynamic allostery' can be transferred to any enzyme that is likely to rely on subtle short time-scale changes in dynamics and network of

non-covalent interactions between residues for activation, regulation and complex formation.

CONCLUSION

Allostery is the process by which binding of a ligand at one site of a protein leads to a functional response at a distant site. Despite its importance in enzymatic regulation, little is known about how the allosteric signal is communicated from one site to the other, especially in the absence of large conformational changes in the protein. Here we have shown that monitoring the pK_a variations of ionizable residues over the course of molecular dynamics simulations in the presence and absence of allosteric regulator can unravel the allosteric communication pathways in proteins. We have verified the predictions of this approach experimentally, by showing that subtle pH variations can mimic the allosteric control, by changing the protonation state of key residues without affecting the catalytic activity of the enzyme or its ability to bind an allosteric effector molecule.

METHODS

Crystallization of *Nme*DAH7PS. *Nme*DAH7PS was expressed and purified as previously described.[21] The resulting protein solution was concentrated to approximately 10 mg mL^{-1}. 1 µL of enzyme solution (9–11 mg mL^{-1}) was mixed with 1 µL of crystallization buffer containing 0.1 M Tris HCl (pH 7.3), 0.2 M trimethyl-amino-N-oxide (TMAO), 600 µM $MnSO_4$ and 15–20 % (w/v) PEG 2000MME. Crystals were grown by hanging drop vapor diffusion over 500 µL of crystallization buffer and the crystallization trays were incubated at 20 ºC. Crystals began to form in 48 hours and were fully formed within seven days. The grown crystals were added to another drop containing 50% SEC buffer containing 200 µM PEP, 100 mM KCl and 10 mM BTP and 50% crystallization buffer also containing either 10 mM Phe and left to soak for 2-6 hours. Crystals were flash frozen using liquid nitrogen in a cryoprotectant solution containing ligand present reservoir solution and 20 % (v/v) PEG400.

Crystallography, Structure determination and refinement. An X-ray diffraction dataset was collected at the Australian synchrotron using the MX1 beamline.[43] The datasets were integrated and processed using XDS and aimless.[44,45] Appropriate cut-off resolution was determined via $CC^{1/2}$≥0.5 ensuring the data was complete in the highest resolution shell.[46] Space group and unit cell parameters for *Nme*DAH7PS with the soaked ligands were the same as those previously identified (PDB code 4HSN)[21] meaning that initial phases could be obtained via molecular replacement using the original structure as a search model in Phaser MR.[47] All ligands and waters were removed from the search model (PDB code 4HSN) before molecular replacement was carried out. Refmac5 was used to generate the electron density map and this was manually analyzed and refined in COOT.[48,49] The quality of the model was optimized by consecutive model building in COOT and refinement with Refmac5. Water molecules were added manually via interpretation of the $|2F_o-F_c|$ map ensuring that they had the ability to hydrogen bond to at least one acceptor or donor. Residues Met1-Asp9 were absent from the model (PDB code 4HSN) used for refinement, however this region of the protein becomes ordered in the presence of Phe and only the first 3 residues are unresolved. Molprobity was used to assess structure quality during refinement cycles and before deposition.[50]

Isothermal titration calorimetry. Binding of *Nme*DAH7PS to L-phenylalanine at pH 7 and 7.8 was measured by ITC using a VP- ITC unit operating at 298 K (Microcal, GE Health-care). Before use, the protein was buffer exchanged against binding buffer [0.5 mM $MnSO_4$ in 50 mM BTP buffer (pH 7 and 7.8)] and all solutions were degassed in a vacuum. Protein concentration was measured by UV absorption immediately before titrations were started. The titrations were comprised of 28 injections, one 2 µL injection followed by 27 10 µL injections of L-phenylalanine. The initial datum point was routinely deleted to allow for diffusion of ligand across the needle tip during the equilibration period. A heat of dilution experiment was measured independently and subtracted from the integrated data before curve fitting in Origin 7.0. For the binding of NmeDAH7PS to L-phenylalanine, 15 µM of *Nme*DAH7PS was used and the syringe contained 1.2 mM L-phenylalanine; and the data were fitted with the two-site sequential-binding model supplied by Micro-Cal.

pH dependent kinetics. The assay system for *Nme*DAH7PS was a modified form of the assay used by Schoner and Herrmann as previously described.[51] The assays to determine the optimum pH of *Nme*DAH7PS activity contained PEP (100 µM) and $MnSO_4$ (100 µM) in 50 mM BTP buffer, 10 µM ethylenediaminetetraacetic acid (EDTA) and E4P (500 µM) which were the correct pH (6.8, 7.3, 7.8 and 8.3) at 298K. Outside of this range, the enzyme dramatically lost activity. The buffer was made up in ultrapure water that had been treated with Chelex resin. After incubation at the required temperature for 5 minutes, the reaction was initiated by the addition of enzyme (2 µL, 1.0 mg/mL). For the pH dependent inhibition kinetics, the buffer contained the same components as above, except 300 µM L-phenylalanine was added before incubation. All assay experiments were performed in triplicate.

Confirmation that substrate binding was not severely affected by pH range. For *Nme*DAH7PS at pH 8.3, the reactions to determine K_m^{E4P} used 100 µM PEP and 32 µM to 375 µM E4P, and for K_m^{PEP} used 500 µM E4P and 3 µM to 100 µM PEP. For *Nme*DAH7PS at pH 6.8, the reactions to determine K_m^{E4P} used 100 µM PEP and 12.5 µM to 250 µM E4P, and for K_m^{PEP} used 500 µM E4P and 7.5 µM to 120 µM PEP. The buffers contained $MnSO_4$ (100 µM) in

50 mM BTP buffer and 10 μM ethylenediaminetetraacetic acid (EDTA). They were initiated by the addition of NmeDAH7PS (2 μL, 1.0 mg/mL).

System preparation for the MD simulations. The crystal structures of *Nme*DAH7PS in the absence (PDB ID: 4HSN) and presence of allosteric Phe (PDB ID: 4UC5) were used as a starting point for the calculations. Missing residues and side chains were reconstructed with MODELLER.[52] The top two tetrameric models generated for each structure, selected on the basis of the lowest MODELLER objective function value and the lowest global DOPE score, were chosen as starting point for the simulations (two simulations for each apo and Phe-bound forms). PROPKA3.1[24] was used to assess the pK_a of ionizable residues and the value averaged by chain and model. The protonation state of these residues for the apo system and for the Phe-bound system was then selected based on these pK_a to simulate a pH of 7.4. Proton position of histidine residues was chosen with MCCE.[39]

MD simulations. All MD simulations were carried out with NAMD[53] using CHARMM all-atom parameter set 22 with the CMAP correction.[54,55] The simulations were run on 1048 cores of the IBM BlueGene/P part of the high performance computing service facility at the University of Canterbury. The particle mesh Ewald (PME) method was used to calculate electrostatic interactions and Van der Waals cut-off was set up at 12 Å. Each system was solvated in an explicit TIP3P water box of 118 Å x 118 Å x 130 Å and the net charge neutralized with Na^+ and Cl^- ions added with a minimum distance of 5 Å from the enzyme and from each other.Using the conjugate gradient energy minimization method, each solvated protein was minimized for a total of 100000 steps first by allowing only water molecules, ions and hydrogens atoms to move (20000 steps), then keeping only the backbone atoms fixed (30000 steps) and finally releasing all atoms (50000 steps). Each system was then gradually headed from 0 K to 310 K over 124000 steps keeping a harmonic restraint of 10 kcal/mol on backbone atoms. The Nosé-Hoover Langevin piston pressure control was then used to simulate the NPT ensemble at 1 atm, 310 K. Each system was equilibrated for 100000 steps with the restraint decreased to 5 kcal/mol, for another 100000 steps with the restraint at 1 kcal/mol, and finally for 1 ns without any restraints. The four equilibrated systems were then each subjected to 200 ns of NPT molecular dynamics simulation at 1 atm and 310 K with the trajectory saved every 10 ps.

Trajectory analysis. The resulting trajectories were aligned to remove any undesired rotational and translational movements of the enzyme. The first 80 ns of each simulation were removed to allow for full equilibration of the systems. Since it has been shown that the dimeric unit is the smallest biological unit capable of both catalysis and allosteric regulation [22], so each dimer of the tetramer were considered as two different trajectories for the purpose of position and distance analysis and combined to effectively lead to 480 ns of fully equilibrated trajectory for each system. After alignment of the resulting trajectories, root mean square deviations (RMSD) and root mean square fluctuations (RMSF) were calculated using Gromacs 4.5 [56]. Correlated motions based on Pearson coefficient were determined with Carma [57]. Distance distributions were calculated in VMD [58] using in-house Tcl scripts using the hydroxyl oxygen of Ser, Tyr, Thr, the center of the two carboxyl oxygens of Glu and Asp, the nitrogen of the ε-amino group of Lys, the center of the guanidinium ion of Arg and the center of the two nitrogens of the imidazole ring of His.

pK_a analysis. For each frame of the tetramer trajectories, the coordinates of the protein without hydrogen atoms, in complex with its ligands (manganese ions and, should it be present, allosteric Phe) were saved as a PDB format. The resulting files were converted from CHARMM PDB format to a more standard PDB format. For each system, PROPKA3.1 was run on each of the 24,000 structure files in parallel using asynchronous dynamic load balancing and the output files parsed in usable data. In order to take full advantages of the different conformational space sampled by each monomeric unit composing the tetramer, the values of interest (pKa, desolvation effect) and each type of interactions (coulombic, side-chain hydrogen bonds and backbone hydrogen bonds) were analyzed in such way that averages value can be obtained (i.e. does not depends on the chain the ionizable residue belongs to). Parsing and analysis of the data were performed using dedicated python scripts. The method and the associated scripts can be easily be transferred to the analysis of other systems and is available upon request.

ASSOCIATED CONTENT

Supporting Information. Tables S1 – S3 and Figures S1 – S7. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* Department of Chemistry, University of Canterbury, Private Bag 4800, Christchurch, Telephone (+64) 3 364 5682, Fax (+64) 3 364 2110, Email: emily.parker@canterbury.ac.nz

Author Contributions

‡These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This work was funded by the New Zealand Marsden fund (UOC1105). Part of this work is funded by UC HPC, the high performance e-research computing service facility of the University of Canterbury. E.J.M.L. is the recipient of a UC BlueFern HPC PhD Scholarship. L.C.H. is the recipient of a

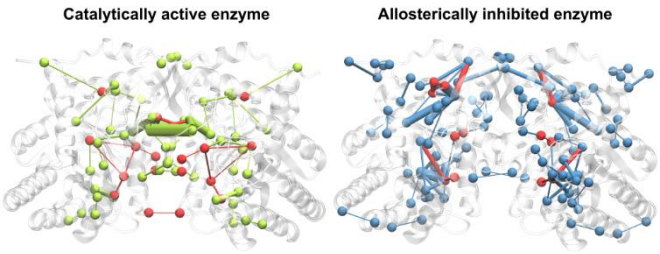
Biomolecular Interaction Centre PhD scholarship. Part of this research was undertaken on the MX1 and MX2 beamlines at the Australian Synchrotron, Victoria, Australia, with usage administered by the NZSG.
The authors would like to acknowledge Dr Jane Allison for stimulating discussions and Dr Sung Bae for his help in writing some of the python scripts used in this work.

REFERENCES

(1) Monod, J.; Wyman, J.; Changeux, J. P. *J. Mol. Biol.* **1965**, *12*, 88.
(2) Koshland, D. E., Jr.; Nemethy, G.; Filmer, D. *Biochemistry* **1966**, *5*, 365.
(3) Cooper, A.; Dryden, D. T. F. *Eur. Biophys. J.* **1984**, *11*, 103.
(4) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. *Nature* **2014**, *508*, 331.
(5) Kern, D.; Zuiderweg, E. R. *Curr. Opin. Struct. Biol.* **2003**, *13*, 748.
(6) Cui, Q.; Karplus, M. *Protein Sci.* **2008**, *17*, 1295.
(7) Goodey, N. M.; Benkovic, S. J. *Nat. Chem. Biol.* **2008**, *4*, 474.
(8) Tsai, C. J.; Del Sol, A.; Nussinov, R. *Mol. Biosyst.* **2009**, *5*, 207.
(9) Smock, R. G.; Gierasch, L. M. *Science* **2009**, *324*, 198.
(10) Feher, V. A.; Durrant, J. D.; Van Wart, A. T.; Amaro, R. E. *Curr. Opin. Struct. Biol.* **2014**, *25*, 98.
(11) Johnson, Q. R.; Lindsay, R. J.; Nellas, R. B.; Fernandez, E. J.; Shen, T. *Biochemistry* **2015**, *54*, 1534.
(12) Sethi, A.; Eargle, J.; Black, A. A.; Luthey-Schulten, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6620.
(13) Vanwart, A. T.; Eargle, J.; Luthey-Schulten, Z.; Amaro, R. E. *J. Chem. Theory Comput.* **2012**, *8*, 2949.
(14) Palmai, Z.; Seifert, C.; Grater, F.; Balog, E. *PLoS Comput. Biol.* **2014**, *10*, e1003444.
(15) Zhou, J.; Bronowska, A.; Le Coq, J.; Lietha, D.; Grater, F. *Biophys. J.* **2015**, *108*, 698.
(16) Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3499.
(17) Suel, G. M.; Lockless, S. W.; Wall, M. A.; Ranganathan, R. *Nat. Struct. Biol.* **2003**, *10*, 59.
(18) McClendon, C. L.; Friedland, G.; Mobley, D. L.; Amirkhani, H.; Jacobson, M. P. *J. Chem. Theory Comput.* **2009**, *5*, 2486.
(19) Lange, O. F.; Grubmuller, H. *Proteins* **2006**, *62*, 1053.
(20) Pace, C. N.; Grimsley, G. R.; Scholtz, J. M. *J. Biol. Chem.* **2009**, *284*, 13285.
(21) Cross, P. J.; Pietersma, A. L.; Allison, T. M.; Wilson-Coutts, S. M.; Cochrane, F. C.; Parker, E. J. *Protein Sci.* **2013**, *22*, 1087.
(22) Cross, P. J.; Heyes, L. C.; Zhang, S.; Nazmi, A. R.; Parker, E. J. *PLoS One* **2015**.
(23) Heyes, L. C.; Reichau, S.; Cross, P. J.; Jameson, G. B.; Parker, E. J. *Bioorg. Chem.* **2014**, *57*, 242.
(24) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525.
(25) Zhou, H. X.; Vijayakumar, M. *J. Mol. Biol.* **1997**, *267*, 1002.
(26) Soares, T. A.; Lins, R. D.; Straatsma, T. P.; Briggs, J. M. *Biopolymers* **2002**, *65*, 313.
(27) Sandberg, L.; Edholm, O. *Biophys. Chem.* **1997**, *65*, 189.
(28) van Vlijmen, H. W.; Schaefer, M.; Karplus, M. *Proteins* **1998**, *33*, 145.
(29) Koumanov, A.; Karshikoff, A.; Friis, E. P.; Borchert, T. V. *J. Phys. Chem. B.* **2001**, *105*, 9339.
(30) Livesay, D. R.; Jacobs, D. J.; Kanjanapangka, J.; Chea, E.; Cortez, H.; Garcia, J.; Kidd, P.; Marquez, M. P.; Pande, S.; Yang, D. *J. Chem. Theory Comput.* **2006**, *2*, 927.
(31) Whalen, K. L.; Tussey, K. B.; Blanke, S. R.; Spies, M. A. *J. Phys. Chem. B* **2011**, *115*, 3416.
(32) Costa, M. G. S.; Batista, P. R.; Shida, C. S.; Robert, C. H.; Bisch, P. M.; Pascutti, P. G. *BMC Genomics* **2010**, *11*, S5.
(33) Baptista, A. M.; Teixeira, V. H.; Soares, C. M. *The Journal of Chemical Physics* **2002**, *117*, 4184.
(34) Lee, M. S.; Salsbury, F. R.; Brooks, C. L. *Proteins: Structure, Function, and Bioinformatics* **2004**, *56*, 738.
(35) Meng, Y.; Roitberg, A. E. *J. Chem. Theory Comput.* **2010**, *6*, 1401.
(36) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219.
(37) Kuhn, B.; Kollman, P. A.; Stahl, M. *J. Comput. Chem.* **2004**, *25*, 1865.
(38) Bashford, D. In *Scientific Computing in Object-Oriented Parallel Environments*; Ishikawa, Y., Oldehoeft, R., Reynders, J. W., Tholburn, M., Eds.; Springer Berlin Heidelberg: 1997; Vol. 1343, p 233.
(39) Song, Y.; Mao, J.; Gunner, M. R. *J. Comput. Chem.* **2009**, *30*, 2231.
(40) Tynan-Connolly, B. M.; Nielsen, J. E. *Nucleic Acids Res.* **2006**, *34*, W48.
(41) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. *Comput. Phys. Commun.* **1995**, *91*, 57.
(42) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. *Nucleic Acids Res.* **2012**, *40*, W537.
(43) McPhillips, T. M.; McPhillips, S. E.; Chiu, H.-J.; Cohen, A. E.; Deacon, A. M.; Ellis, P. J.; Garman, E.; Gonzalez, A.; Sauter, N. K.; Phizackerley, R. P.; Soltis, S. M.; Kuhn, P. *Journal of Synchrotron Radiation* **2002**, *9*, 401.
(44) Evans, P. *Acta Crystallographica Section D* **2006**, *62*, 72.
(45) Evans, P. *Acta Crystallographica Section D* **2011**, *67*, 282.
(46) Karplus, P. A.; Diederichs, K. *Science* **2012**, *336*, 1030.
(47) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658.
(48) Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallographica Section D* **2011**, *67*, 355.
(49) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallographica Section D* **1997**, *53*, 240.
(50) Chen, V. B.; Arendall, W. B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallographica Section D* **2010**, *66*, 12.
(51) Schoner, R.; Herrmann, K. M. *J. Biol. Chem.* **1976**, *251*, 5440.
(52) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
(53) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
(54) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.;

Lau, F. T.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
(55) Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd *J. Comput. Chem.* **2004**, *25*, 1400.
(56) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29*, 845.
(57) Glykos, N. M. *J. Comput. Chem.* **2006**, *27*, 1765.
(58) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

Table of Contents artwork



Supporting information

Table S1. Data Collection and refinement

	*Nme*DAH7PS
	L-Phenylalanine
Data Collection	
Crystal system; space group	Monoclinic, $P12_11$
Unit cell parameters	
a, *b*, *c* (Å)	73.61, 143.50, 75.18
α, β, γ (º)	90, 96.18, 90
Resolution range (Å)	47.82-2.19 (2.23-2.19)
Measurements	301997
Unique reflections	79522
Redundancy	3.8
Completeness (%)	99.9 (100.0)
I/σ (I)	9.3 (1.2)
R_{merge}	0.084
$CC_{1/2}$	0.50
Wilson B value $(Å)^2$	39.81
Matthews coefficient	2.60
Refinement	
R_{work}	0.1890
R_{free}	0.2357
Chain length	351
Observed number of residues	345 (All Chains)
Water molecules	330
Other (Mn^{2+}, PEG, SO_4^-)	9
Ligand	4
Mean B $(Å)^2$	
Protein	48.59
Water	43.82
Other	67.46
Phe	37.13
R.m.s.d from target values	
Bond lengths (Å)	0.0095
Bond angles	1.2712
Dihedral angles	0.0690
Ramachandran	
Preferred (%)	97.69
Allowed (%)	1.73
Outliers (%)	0.58
PDB Entry	4UC5

Table S2. Calculated Phe dissociation constant from ITC experiments

	Site 1	Site 2	Model
pH 7	**3.3 ± 0.8 μM**	**11 ± 1.3 μM**	**Sequential two site**
pH7.8	**2.0 ± 0.7 μM**	**8.6 ± 1.2 μM**	**Sequential two site**

Table S3. Substrate affinities at pH 6.8 and 8.3

	pH 6.8	pH 8.3
K_m^{E4P} (μM)	36.7 ± 1.8	106 ± 15
K_m^{PEP} (μM)	15.4 ± 0.9	7.0 ± 0.6

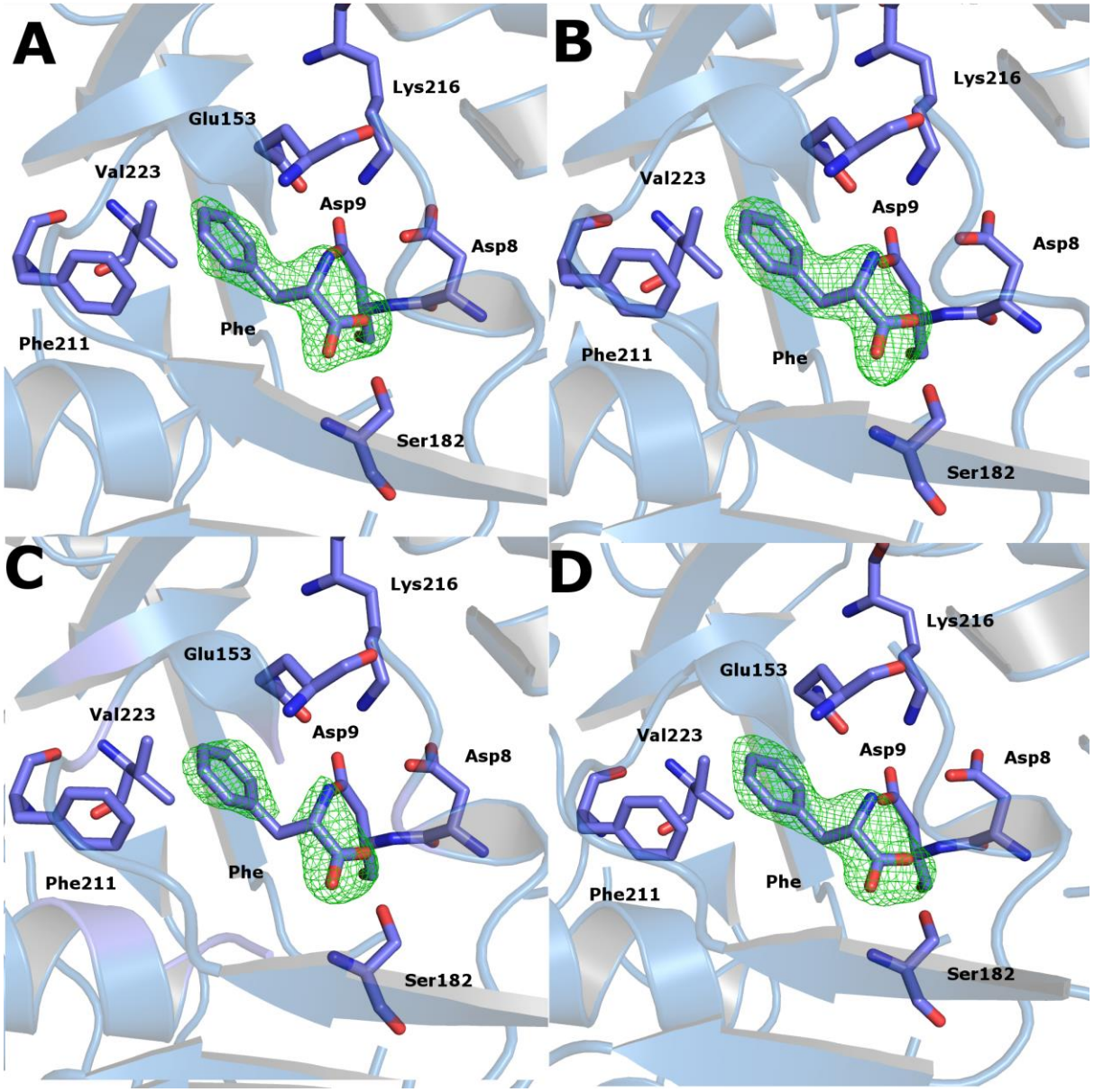


Figure S1. $|F_o\text{-}F_c|$ maps for L-phenylalanine bound at the allosteric site contoured at 3σ. The four chains of NmeDAH7PS are indicated. Top left: chain A, op right; chain B, bottom left: chain C and bottom right: chain D.

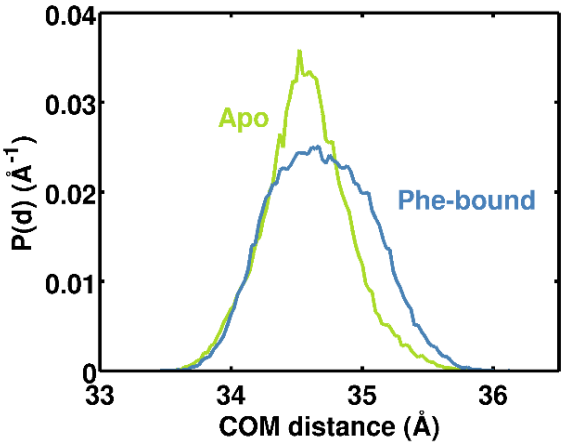


Figure S2. Distribution of the distance between the centre of mass (COM) of each monomer for the apo (green) and Phe-bound (blue) MD simulations.

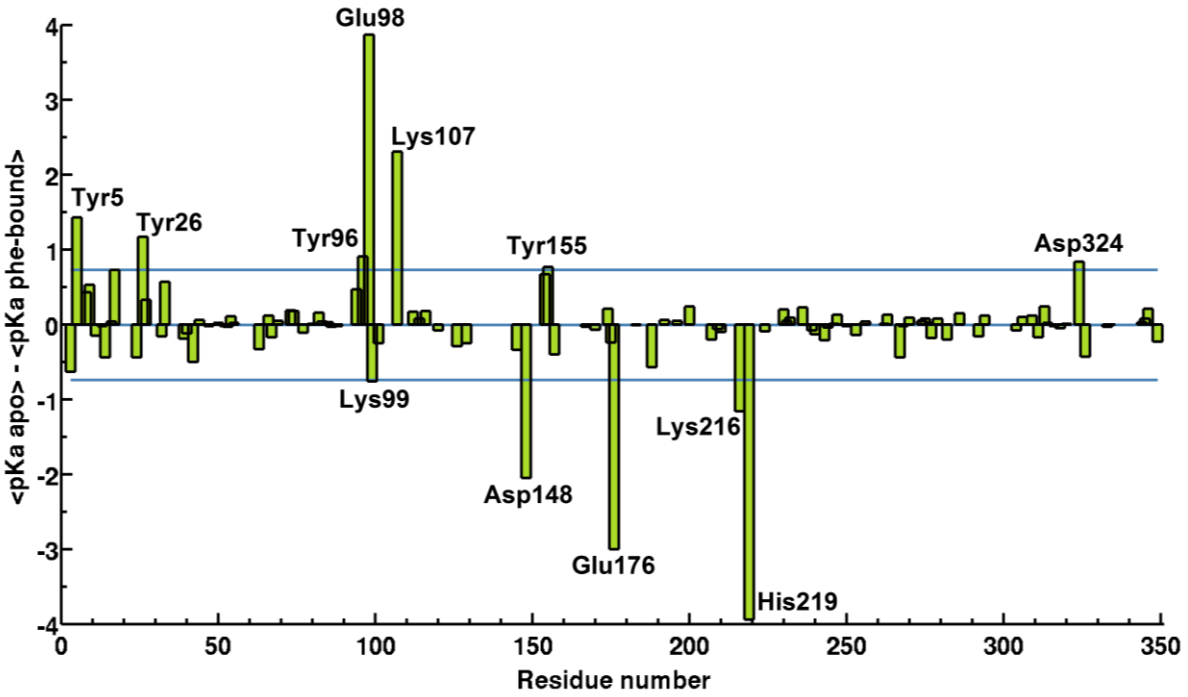


Figure S3. Calculated pK_a difference of ionizable residues between apo and Phe-bound forms of *Nme*DAH7PS, averaged over the course of the MD simulations. The middle blue line represents the average difference in pK_a between the apo and Phe-bound form and the top and bottom blue lines correspond to the average ± one standard deviation. Residues that present a major pK_a shift and therefore might be involved in the allosteric communication pathway have been selected on the basis of their distance from the average ± one standard deviation.

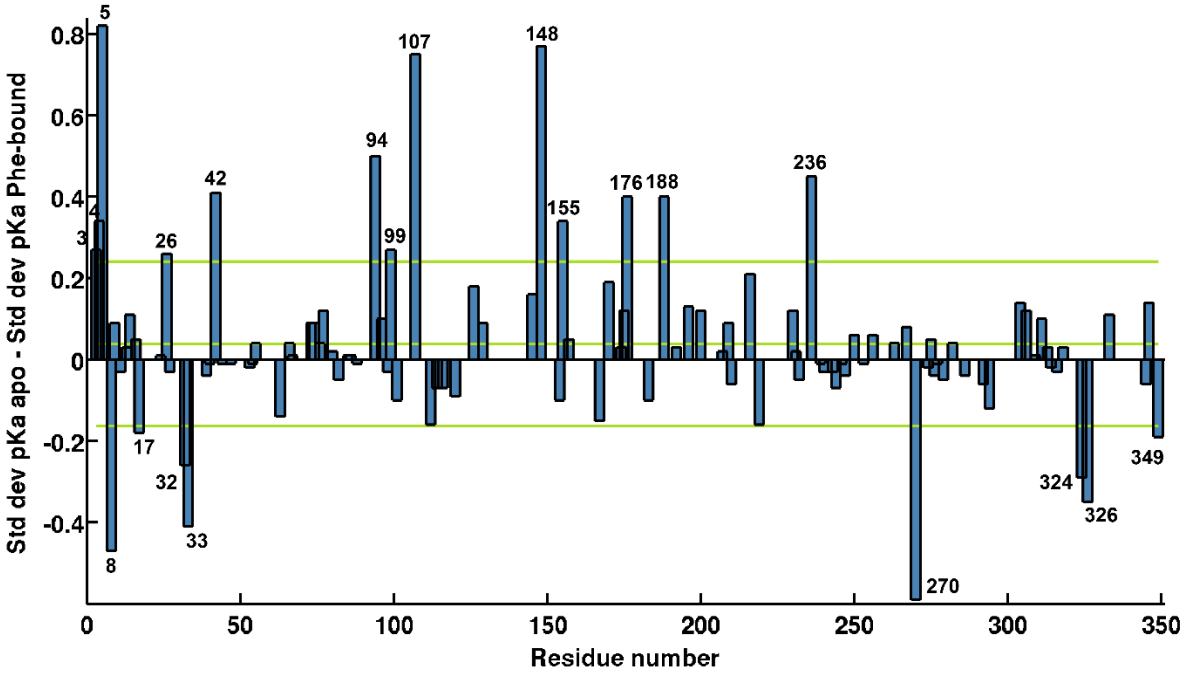


Figure S4. Standard deviation of the pK_a of the ionizable residues. Difference between the standard deviation of the pK_a of the ionizable residues of the apo and Phe-bound forms, over the course of the MD simulations. The middle green line represent the average difference and the upper and lower green lines correspond to the average ± one standard deviation. Residues that might be implied in the allosteric communication pathway have been selected on the basis of their distance from the average ± one standard deviation.

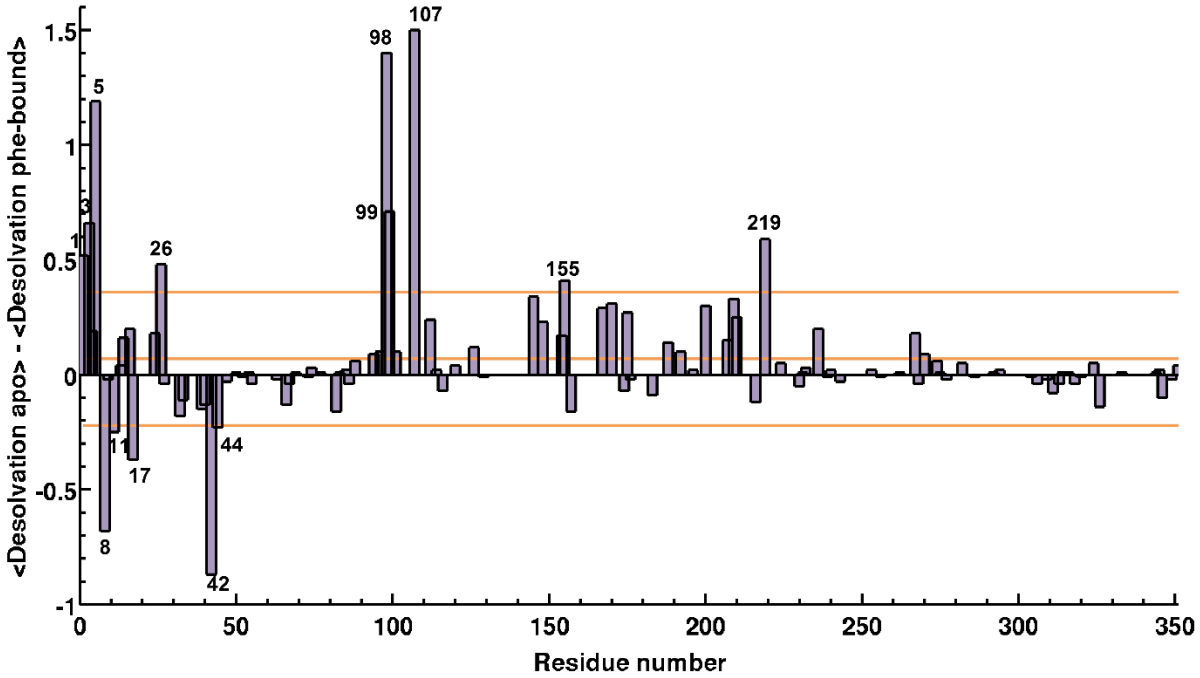


Figure S5. Average desolvation values of ionizable residues. Difference between the desolvation values of the ionizable residues of the apo and Phe-bound forms, averaged over the course of the MD simulations. The desolvation is given in unit of pK_a. The larger the value, the more buried the residue is. The middle orange line

represent the average difference and the two other orange lines correspond to the average ± one standard deviation.

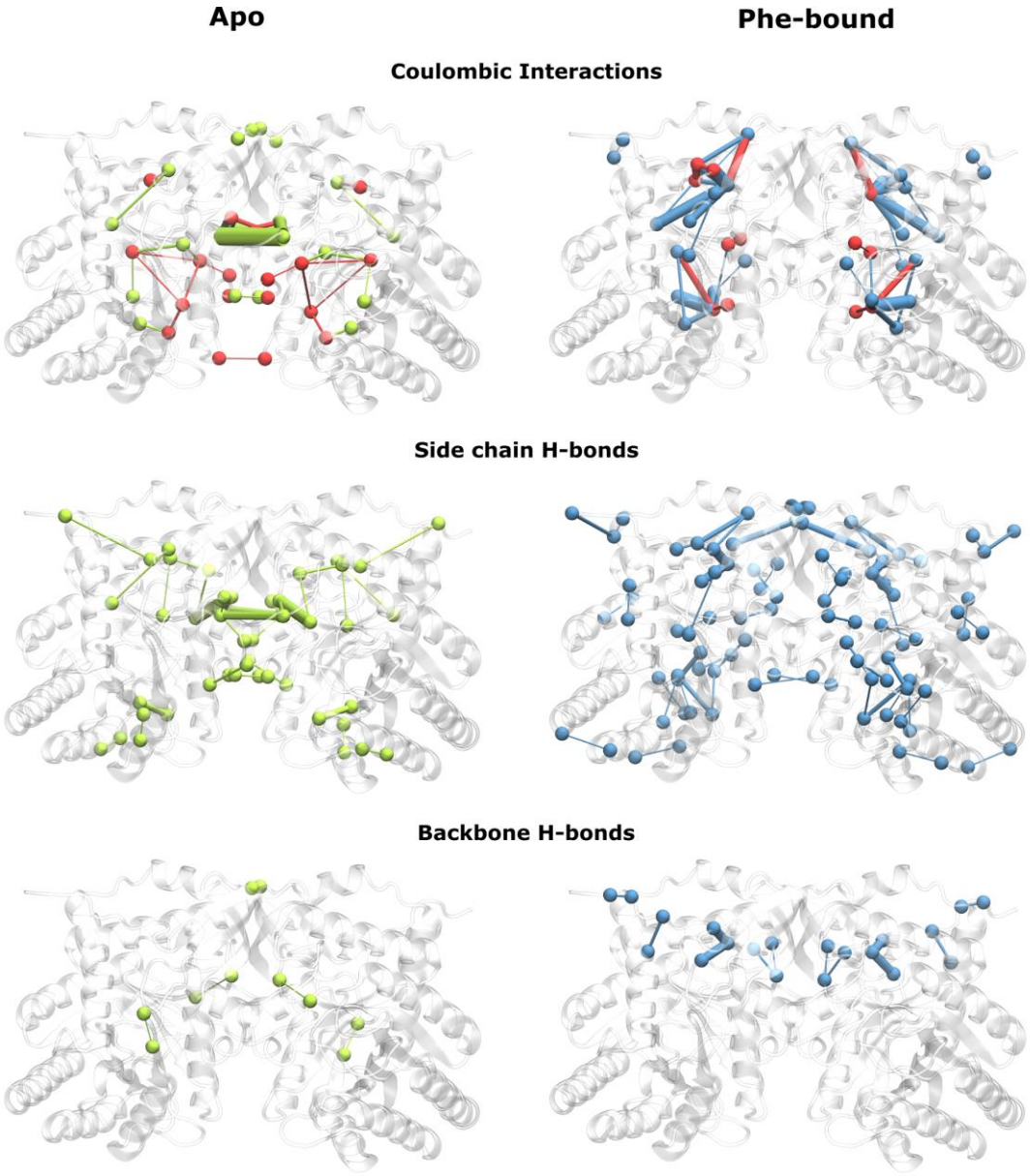


Figure S6. Detailed differential maps of interactions between the apo and Phe-bound forms. Details of the coulombic interactions, side chain H-bounds (between the ionizable group of a residue and the hydrogen donor or acceptor function present on the side chain of another residue) and backbone H-bonds (between the ionizable group of a residue and the backbone of another) that are only present or stronger in the apo (left column) or the Phe-bound (right columns) trajectories. Nodes represent interacting residues and the edges between nodes indicate the interactions, with the diameter of the edge accounting for the strength of the interaction. For the apo

form, favorable interactions are in blue and unfavorable ones are in orange, and for the Phe-bound form, favorable interactions are in red and unfavorable ones are in purple.

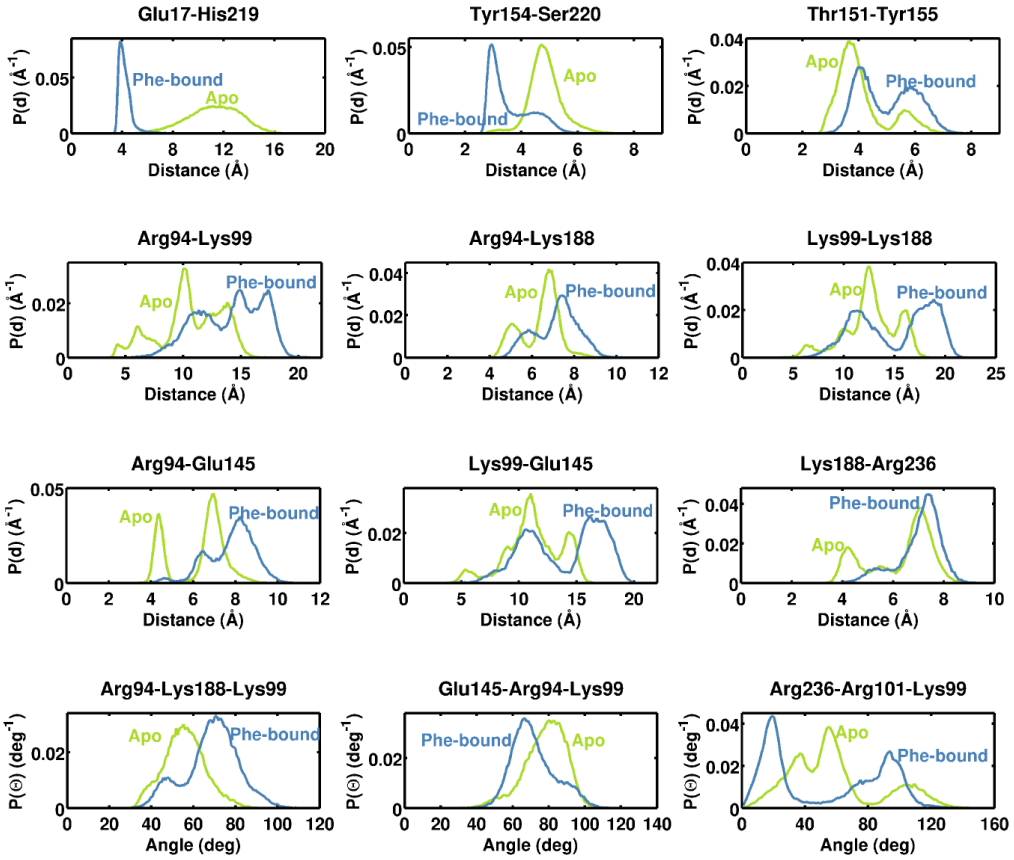


Figure S7. Distribution of the distances between ionizable groups of interest. Frequency distribution of the distances for the apo (green) and Phe-bound (blue) forms. The distances are calculated using the hydroxyl oxygen of Ser, Tyr, Thr the center of the two carboxyl oxygens of Glu and Asp, the nitrogen of the ε-amino group of Lys, the center of the guanidinium ion of Arg and the center of the two nitrogens of the imidazole ring of His.